

What price independence?

The British government has scraped together a little money to save its universities from catastrophe, winning in return a series of promises that academics may not be able or willing to keep.

BRITISH universities cannot afford to decline last week's offer of extra funds (see p. 97), but they will pay a high price for only a little extra. Not that the strings attached to the offer by Mr Kenneth Baker, the new Secretary of State for Education and Science, are all that onerous. He demands selectivity in research, the elimination of small departments, better management, the abolition of academic tenure, more technology transfer to industry, more explicit concern about academic standards and more help with the supply of mathematics and physics teachers in the schools. What could be more reasonable? Unfortunately, there is a catch.

What Mr Baker asks is not new. Some of his requests are even matters on which the university system itself has taken the initiative. Selectivity, for example, has been a slogan for at least twenty years. More recently, it has been the practical response of the University Grants Committee (UGC) to the shrinkage of its budget; its intention to concentrate part of its funds on institutions with a reputation for research is two years old, but began painfully to bite with this academic year. Similarly, for better or worse, British universities are now thoroughly steeped in the doctrines on management put out last year by what is called (after the chairman of a committee) the Jarratt report, commissioned jointly by the Committee of Vice-Chancellors and Principals and UGC. So why fear the gift of extra funds when the strings attached require largely the pursuit of present goals?

The obvious danger is that the prudent intentions of titularly autonomous institutions are turned into external obligations that can (and will) be monitored. Selectivity? A single civil servant in his spare time will be able to inspect the disparities in the treatment by UGC of individual universities and complain if they are too small. Better management? UGC has undertaken to carry out the first invigilation itself. Academic standards? The vice-chancellors, after what has evidently been a busy summer, have put out a mixed dossier of codes of practice (on matters such as the training of academic staff and the role of external examiners) and of half-considered means of assessing academic performance (citation counts, for example), giving further hostages to the external monitors. The vice-chancellors have even promised to produce a report on present practices of awarding academic tenure; without promising to abolish tenure, the vice-chancellors' letter promises to "aim" at ensuring that people "are not initially appointed with tenure in its strongest form where this can be avoided".

Accountability

But should not universities supported by public funds be publicly accountable? Of course. The danger in the new relationship between the British government and its universities is that the social and economic value of the academic system will be judged by criteria devised for managerial purposes only. It should be especially galling for academics, and alarming as well, that Baker was able to attach to last week's written statement copies of letters from the chairmen of UGC and of the vice-chancellors' committee embodying written promises that they will conduct themselves appropriately in the future.

How could that have come about? In the idiom of British public life, the circumstances are easily reconstructed. Baker differs from Sir Keith Joseph, his predecessor, in preferring to make friends rather than enemies. He is also adroit, so much so that the failure of his first project for the present government, the reinvigoration of the British economy by the encouragement of cable television, has left his reputation unscarred. At some early meeting with the chairmen of UGC and the vice-chancellors' committee, he would have explained that he personally is sympathetic to their difficulties (and appalled at the prospect that one or more universities might have to close), would have added that the government in which he shares collective responsibility is unfortunately convinced that academics are an expensive bunch of unproductive layabouts, would have listened sympathetically to the universities' account of how strenuously they have tried to put their own houses in order (and would generously have advised them to spend more on making their efforts better known), would have offered to fight hard for a little extra cash in the run-up to the annual settlement (last week) of next year's public spending and, pleading that he himself would be occupied trying to bring the semblance of industrial peace to British schools, would have asked that academics should help him and their own cause by providing the ammunition he would need.

Contrition

The sting has worked. There will be a little extra money. Baker's civil servants would have provided guidance (in the British idiom) about the kind of ammunition they would need (but their first proposal for a tripartite statement appears to have foundered, improbably at a gathering of academics in Penang). But Baker has been able to deliver to his colleagues a prize that eluded his predecessor — a show of contrition for imagined misdemeanours by the universities, a promise to do better in future and a set of measuring-rods with which they may be beaten if the occasion arise.

Nothing that happened last week touches the true problem of the British university system, which is structural and neither financial nor managerial. The universities boast that they are autonomous, but that is less than half the truth. In reality, they are public corporations almost totally dependent on public funds (on UGC and on local education authorities for their teaching functions, on the research councils for research). In the halcyon 1960s, when the only problem was that of investing quickly enough in higher education, UGC's performance easily upheld the British conceit that it is a uniquely judicious way of channelling public funds towards private ends.

The novelty, now, is that British governments, having become by accident providers of higher education, have also for some time become disenchanted with the product (or at least persuaded that higher education can be had more cheaply in other kinds of institutions than universities). UGC, no longer able to act the benefactor, has taken to telling particular institutions that they cannot teach this, or that ... or else. One of UGC's latest projects, an enquiry into the viability of "small" departments of physics and chemistry, will almost certainly discover



that they are financially uneconomic and will recommend that they should be abolished or amalgamated. UGC will regret, like the institutions concerned, that universities thus deprived will hardly meet the classical concept of the university.

Especially under its present enlightened management, there is no reason to believe that UGC wants to become what it is well on the way to being — every British academic's dean's office. Nor can Baker seek the kind of centralization now being forced on the system by financial stringency. He should know that there is a respectable way out, one that would get the government out of the business of setting the specific goals by which universities conduct themselves. He would be nudged in that direction if the committee now looking at UGC's own function, due to report in the next twelve months, recommends regretfully that UGC should be abolished.

But how could British universities survive without UGC? By being challenged with the offer of the autonomy of which they boast. It is now five years since they were told (by UGC) that they would be penalized financially if they exceeded certain student numbers; the threat of penalty has been lifted, but target student numbers remain the basis on which UGC allocations are calculated. One result is that successful universities cannot grow as quickly as they might, while less successful institutions are kept in being, at least in the short run, by continuing subventions. The prospect, as UGC enforces rationalization for sound economic reasons, is that many of these institutions will become lop-sided places, covering only part of the ground of scholarship and thus even less suitable for, and probably even less attractive to, potential students. The truth is that British universities are, on the average, too small for the smaller among them to be complete. Their size probably makes them uneconomic as well. The general interest would be served by amalgamation, for which, under present arrangements, there is no incentive.

But how could autonomy be afforded? Curiously, the financial problem is not nearly as insurmountable as it may seem. A third of the UGC budget is reckoned to cover research support, leaving £1,000 million a year in round numbers for the support of other academic functions in universities. Public funds are channelled to institutions of higher education in the form of tuition fees fixed far below average or even marginal cost. Although local authorities administer the system, central government meets the cost. So why not allow tuition fees to rise again (they were arbitrarily cut four years ago when Baker's predecessor feared they would partly protect some universities from the penalties for quota-breaking), meeting the cost from the central budget of UGC? If managed over a predetermined period of time, and coupled with the understanding that students would choose between universities, the result would be both autonomy and a measure of competition between institutions at present persuaded that they must all sink together.

But, Baker's civil servants will ask, what assurance will there then be that the universities are publicly accountable? How will it be possible to tell that academics are indeed carrying out a modicum of research, or that they are taking full account of citation counts in promoting their own members to more senior posts? These questions are of course irrelevant. If the universities were financially autonomous, dependent on public funds only indirectly through the preferences of students for courses of study and for places at which to follow them, the universities would have an incentive which they at present lack to manage their affairs efficiently. They would also have an incentive towards diversity, of which Britain has too little. The endless quarrel with the government about tenure as universities came to put their institutional survival before the interests of individuals would end. Some might even discover that they need not stick slavishly to UGC "norms" for the number of students per member of teaching staff any more than they are compelled to follow UGC standards for the equipment of buildings. To be sure, in such a regime, only the better universities would prosper. Is there anything wrong with that? □

Research in trouble

Mr Baker's little extra has also created a dilemma in which leaving CERN may be the only choice.

THE research half of the British academic enterprise will be hardly better pleased than the university part with last week's events. Outwardly, the research councils will have an extra £39 million to spend next year, over and above inflation, or more than 6 per cent of total public support for basic research. One immediate snag is that £14 million of the extra offered last week was also offered just under a year ago by Sir Keith Joseph, and has long since been counted on by the Advisory Board for the Research Councils (ABRC) on whose advice the sum available is eventually divided among the five research councils. Another and more serious difficulty is that the British basic research enterprise is yet again in the throes of a crisis thrust upon it by circumstances over which it has no control. Sterling has once again depreciated, with the result that the Science and Engineering Research Council (SERC), which must find from its budget international subscriptions to organizations such as CERN, the European high-energy physics laboratory, is short of roughly £20 million, nearly a quarter of what it spends each year on research grants to academics. ABRC will have a stark choice at its two meetings later this month: either give the whole of Mr Baker's extra money to the research council in most urgent need of it, letting the others smoulder in penury for another year, or invite SERC to face a decision that has been staring it in the face for the past three years, whether to pull out of CERN.

It is a curious state of affairs. Except for 1980, when sterling appreciated against European currencies (and SERC's predecessor spent its windfall gain on extra equipment grants for universities), the crisis is chronic. Nobody is astonished, for sterling is a steadily depreciating currency. To saddle the research councils with this perpetual difficulty is of course unfair; while the international organizations to which SERC subscribes have scientific objectives, the government also has a political interest in their continuation. With its new policy on space research, the British government would not readily agree to a decision of SERC to save its annual subscription to the European Space Agency by pulling out of that. British membership of CERN is a political matter for different reasons; the British government has an interest in seeming no less faithful a European state than its competitors. But, with the exception of a single year when the Treasury was persuaded to do the decent thing, there has been no occasion in the past decade when the research councils have been adequately bailed out. Members of the council's peer review committees meeting in the past few weeks to consider new grant application have been startled to be told that SERC is £20 million short, for which reason it may be necessary to postpone the dates at which some successful applications become effective. Of necessity, the grants postponed will be those that can be put off practically; they may well include the more promising projects, which is no way to run a research enterprise already low on morale.

This is why withdrawal from CERN is yet again a live issue. ABRC's own inquiry under Sir John Kendrew concluded that it would be better, first, to negotiate a reduced budget (or British contribution), which is now the purpose of a study by CERN members due to be completed next year. It would be discourteous to pull out before that review is completed, but SERC may have no choice. The way things have turned out in Britain, with the research community persuaded that the future is only bleak, the decision may not matter much. But if there is a chance of rescuing something from the British enterprise, it is more important to give a few young people with bright ideas a chance to turn them into reality than to continue membership of even such a splendid organization as CERN. But if SERC goes down that road, it should promise itself and its constituents that it will also look hard at the need for all of its own domestic facilities. □

British penury relaxed

Increases for universities just avert disaster

BRITISH universities and the research councils are to benefit from a modest cash injection in the two years from next April. This is their share of a £4,000 million package of spending increases unveiled last week by the government, and widely taken to be part of the preparation for the next general election, which must be held before the spring of 1988.

Although education as a whole receives the lion's share of the whole package, which includes £460 million for the improvement of the pay of teachers in schools, the universities are unimpressed with their share. This amounts to an extra £61 million in 1987-88 and an extra £71 million the following year. The research councils are to get an extra £39 million a year, beginning with the next financial year next April.

The Committee of Vice-Chancellors and Principals (CVCP) welcomes the extra funds as anybody would welcome "water after years of drought", saying that the new budget will do little more than avert disaster. CVCP was especially indignant at the meagreness of the increase for the research councils, saying that the sum is not only disappointing but that it does not compare with the investment in research by Britain's industrial competitors.

Announcing the increases last week, Mr Kenneth Baker, Secretary of State for Education and Science, more than hinted that there are strings attached. In particular, the university increases are dependent on continuing progress in the rationalization of research and the channelling of sufficient funds into applied research.

Baker, who succeeded Sir Keith Joseph as secretary of state only in June, said last week that he "would expect" the University Grants Committee (UGC) to provide for some early retirement and the recruitment of new young academic staff and special initiatives in teacher training to help deal with the shortage of teachers of mathematics, physics and technology in British schools.

The execution of this policy is in the hands of UGC, which began the process of selectivity last May, when universities were told individually that public support

for this academic year would for the first time be linked with research performance. The equity of that unequal distribution of funds was widely disputed (see *Nature* 322, 299; 1986); by the end of next year UGC and the research councils are to produce a joint statement of the criteria that will be followed in the further pursuit of selectivity in research.



Kenneth Baker's cash has strings attached.

Meanwhile, UGC says that whatever the funds available through the research councils, "universities separately need to draw up and carry out plans for the selective distribution of the resources they devote to research". This will lead to "better research through the concentration of effort in strong departments".

British universities are also under continuing pressure to improve their internal management along lines recommended by the Jarratt committee in 1985; responses by individual universities are now in the hands of UGC. The government is also pressing for common criteria of academic quality and the abolition of academic tenure.

UGC is responding to the further pressure to rationalize the pattern of university activity in Britain by a series of reviews of academic areas. Reviews have already been completed of teaching and research in pharmacy, agriculture, oceanography, Italian and Scandinavian studies.

There will be further pressure on the university system later this month, with the planned publication of a Cabinet Office report on publicly supported research. The government has not hidden its discontent with the efficiency with which research is transferred from universities and government research establishments to the commercial sector. **Bill Johnstone**

Meningitis

Puzzling cluster in rural Britain

THE incidence of meningitis in the small English parish of Stonehouse in Gloucestershire, which is 14 times the British average, has now prompted the British Department of Health to mount a special investigation. Forty doctors, nurses and scientists will shortly complete the first phase of a programme to try to determine the reason. It is hoped that the 7,000 residents of Stonehouse will agree to be tested.

The subject of the study is principally a bacterial strain of meningococcal meningitis (B15R) which is resistant to Group B Type 15 sulphonamide. In the past five years, 53 cases of meningitis, mostly B15R, have been diagnosed in the Stroud area of Gloucestershire (population 103,600), which includes Stonehouse. Since 1984, Stonehouse has recorded 10 out of 12 diagnosed cases of the disease from among the 2,000 people living on one housing estate. The national average for the incidence of meningitis is 1 per 100,000 of the population. The incidence in the Gloucester district is five times that figure, and that of Stroud is double again with 9 cases per 100,000.

The research project is being coordinated by Gloucester Public Health Laboratory and the Gloucester Health Authority with support from the Department of Health and Social Security (£57,000) and the Meningitis Trust (£20,000), a charity recently established in the Gloucester area to fight the disease.

The volunteers in the programme will submit themselves to swabs and blood and saliva tests. The results of the first two categories of data are expected to identify the carriers of the disease and those with immunities; it is hoped that the other, conducted by a team from the University of Edinburgh, will indicate the type of person susceptible to contracting it. The Edinburgh team hopes to substantiate its theory that those who do not secrete blood group substances into their saliva are more likely to develop meningitis.

Although the disease can be successfully treated with antibiotics if diagnosed early, some young victims have died within two days. Fever, rash, vomiting and eventual coma are the classic symptoms. But although the under-fives are more susceptible to the ordinary form of the disease, the B15R strain is claiming many of its victims from among young adults, most of them teenagers.

The numbers of diagnosed cases of meningitis peaked in 1974 with 1,296 cases and dropped to an all-time low ten years later with 404. But last year the figure was 549 and 900 cases are expected this year.

Bill Johnstone

University Grants Committee grants (£ million, year starts 1 April)

	1986-87	1987-88	1988-89
Previous plans	1,318 (108)*	1,356 (106)	1,392 (110)
November 1986 plans	—	1,413 (110)	1,463 (110)
Increase over previous plans	—	57 (4)	71 (0)
		Total: 61	Total: 71

* Values are for the recurrent grant, with equipment grant given in parentheses.

US mid-term elections

Power jockeys in new race

Washington

LAST week's elections put the Democratic Party front and centre on the Capitol Hill stage. With control of both the Senate and the House of Representatives, the Democrats will carve up the spoils of their victory in the next few weeks, chiefly the chairmanships of the committees that determine the legislative agenda for the next session of Congress. But a new majority party will hardly be a panacea for budget troubles, and the Republicans will now have the luxury of watching the Democrats make the difficult choices.

The Democrats gained control of the Senate, and now enjoy a 55–45 majority. Several extremely close races delayed final calculations, but the Democrats will add at least 5 seats to their 253–182 majority in the House of Representatives. Ironically, several of the Republicans being replaced as Senate committee chairmen are reckoned more liberal than their Democrat successors. An example is John Stennis, a conservative southern Democrat, who takes control of the powerful Appropriations Committee from Mark Hatfield, a liberal Republican from Oregon.

The chairmanships of the key appropriations subcommittees controlling the purse strings for the National Science Foundation (NSF) the National Institutes of Health (NIH) and other government science agencies are still uncertain, but present chairmen have treated science agencies generously. Lowell Weicker (Republican, Connecticut) has been a strong ally of NIH as chairman of the health subcommittee, and Utah Republican Jake Garn has been favourably disposed toward NASA (the National Aeronautics and Space Administration) as chairman of NASA's subcommittee. Garn was a passenger aboard the space shuttle in 1985.

The Strategic Defense Initiative (SDI), already under attack in Congress, will probably face greater resistance from the Democrats taking over military and energy committees. Both Sam Nunn of Georgia and Louisiana Democrat J. Bennett Johnston have criticized SDI.

A dramatic philosophy shift will occur on the Labor and Human Resources Committee when Orrin Hatch of Utah steps down in favour of Edward Kennedy of Massachusetts. Hatch, a conservative, has been an outspoken opponent of abortion, favouring a constitutional amendment making it illegal. Kennedy is a liberal and a strong proponent of government-supported health-care programmes.

In the House of Representatives, membership of science committees is unlikely to change substantially. The most significant

difference will be the departure of Don Fuqua from the House Science and Technology Committee. Fuqua, who retired this year and will take a job as a lobbyist for the space industry, will be replaced by Robert Roe of New Jersey.

Although science agencies fared relatively well in the most recent budget cycle, next year will probably be a different story. The Gramm–Rudman deficit reduction act will force still larger cuts to reduce the budget deficit.

William Wells of George Washington University, who follows congressional science policy for the American Association for the Advancement of Science, says a harbinger of next year's problems was seen in the House of Representatives appropriation for the NSF research budget. Although NSF has received favourable treatment from the Democrat-controlled House, this year the appropriation for NSF-sponsored research was \$146 million less than the administration request.

With only a fixed amount of money to dispense, the House subcommittee chose to protect certain housing and environmental programmes at the expense of NSF. In the final budget compromise with the Senate some of that money was restored, and overall NSF received about an 8 per cent increase in funds. According to Wells, in the next Congress just holding steady will be a "a remarkable achievement".

Joseph Palca

British SDI cash

THE value of SDI contracts in Britain has not so far approached the reported \$1,500 million target sought by British politicians at the outset, but their total value now exceeds \$14 million. Contracts and sub-contracts include, according to the Federation of American Scientists and the Council on Economic Priorities, Barr and Stroud \$250,000; Cap Scientific \$250,000; British Aerospace \$1.25 million; Ferranti \$508,000; General Electric (UK), \$1.75 million; Hunting Engineering \$1 million; Logica \$700,000; PA Defence Services \$750,000; Plessey \$1.45 million; Racal \$250,000; Scicon \$250,000; Short Brothers \$250,000; Thorn EMI \$500,000; UK Atomic Energy Authority \$4.3 million; UK Ministry of Defence \$1.4 million.

The Science and Engineering Research Council's Rutherford Appleton Laboratory is also now negotiating a \$15 million research contract for the Super-Sprite Laser Facility. In the British Isles those doing SDI work include University College, Cork, Heriot-Watt University and the University of London. Tim Beardsley

Strategic defence

SDI should change course

Washington

RESEARCH for the Strategic Defense Initiative (SDI) continues to be an obstacle to superpower arms negotiations, but a compromise on defensive research that meets the legitimate interests of both sides should be possible, according to the Aspen Strategy Group, an independent think-tank.

The group's study, *The Strategic Defense Initiative and American Security*, published on 22 October, is critical of the current SDI programme, saying its large-scale mockups are inadvisable for technical reasons and erode confidence in the Anti-Ballistic Missile (ABM) Treaty. But the Aspen group says "progress towards a new [strategic arms] agreement should be possible" if the United States modifies its position that all development and testing of space-based defensive systems is allowed, in return for the Soviet Union dropping its demand that only laboratory research is possible.

The co-chairmen of the Aspen study were William J. Perry, a former under-secretary of defence under President Carter, and Lt-Gen. Brent Scowcroft, national security adviser to President Ford. The group includes senior politicians, academics and representatives of industry.

The group sees valuable opportunities to trade arms cuts for a commitment not to give notice of withdrawing from the ABM treaty for 10 years — this would mean forfeiting "very little in technical terms". The study finds "virtually no prospect" of a significant and effective population shield against a responsive enemy this century. And it says the SDI programme's large field experiments "appear to be driven by political considerations" and risk "freezing in" technologies prematurely in what should be a long-term programme.

The Aspen group also says the United States has conducted SDI experiments "whose consistency with [the ABM treaty] is measured by criteria that we would probably reject if the Soviets used them to justify their programmes". The US administration has proposed a new interpretation of the treaty that would allow testing of SDI components as long as they were not actual weapons.

The group recommends that SDI research should be refocused on more long-term "high-leverage" goals, including the problem of mid-course discrimination of warheads from decoys and that defensive research should have a near-term goal of preventing or responding to a Soviet break-out of the ABM treaty and preserving options for selective defences.

Tim Beardsley

Canadian budgets cut

Sacrificing all for space effort

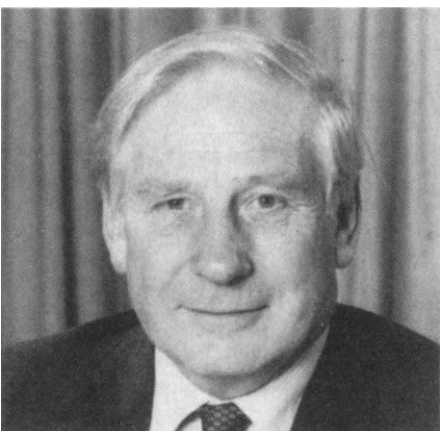
Washington

CANADIAN science, already troubled by an uncertain financial future, took another body blow when the National Research Council (NRC) said recently that it would trim \$20.5 million from its \$400 million budget for the fiscal year starting next April. Particularly painful to those affected is that much of the money being trimmed from basic research efforts is being ploughed into the Canadian space programme, in particular for Canadian participation in the US space station.

The government makes no bones about the fact that some research projects will have to be curtailed or eliminated to support the space effort. In the next five years, NRC expects to realize an additional \$125 million for its space station activities, but in the short term it has had to find \$14.6 million for space activities from its existing budget.

Ralph Nicholls, a space scientist at York University, Ontario, and chairman of the Canadian advisory committee for scientific uses of the space station, says it is unfair to blame NRC for implementing government science policy. Nicholls says Canadian space scientists have also been hurt by the government's decision to eliminate the rocket and balloon programme, and now Canadian scientists find themselves in the same boat as their counterparts in the United States, who are having to wait for space shuttle flights to resume before their scientific projects can start again.

New chairman for vice-chancellors



Sir Mark Richmond to VCVP.

PROFESSOR Sir Mark Richmond, vice-chancellor of the University of Manchester, will become chairman of the UK Committee of Vice-Chancellors and Principals in July 1987 in succession to Mr Maurice Shock, vice-chancellor of the University of Leicester. □

Although the cuts fell relatively evenly among the major divisions of NRC — biological sciences, chemistry, electrical engineering, astrophysics, construction research, mechanical engineering, aeronautics and physics — perhaps no area feels the cuts more keenly than radioastronomy. Canada has two major radioastronomy facilities; the Dominion Radio Astrophysical Observatory at Penticton, British Columbia, and the Algonquin Radio Observatory at Algonquin Park, Ontario. Both are under the control of the Herzberg Institute of Astrophysics, itself supported by NRC.

Institute director Donald Morton says there were plans to refurbish the Algonquin Park observatory to let it operate at shorter wavelengths. But when it became clear that the necessary \$10 million would not be forthcoming, and the institute's budget was being trimmed by \$1.1 million, attention focused on the viability of Algonquin Park. Although Morton insists no decision about the observatory's future will be made until next spring, he admits that prospects are not good.

Paul Wesson, a radioastronomer from the University of Waterloo, believes a decision to close the Algonquin Park Observatory would make Canada a "non-competitor" in international radioastronomy. Even without the refurbishment, the observatory is the largest steerable antenna in North America devoted solely to radioastronomy. The planned improvement would allow the telescope to operate in the 1.5–4 mm range, making it the second largest millimetre-wavelength observatory in the world.

Wesson worries that because radioastronomy does not provide readily appreciated benefits to society, it is particularly vulnerable to future cutbacks.

Optical astronomy has also been hurt by the NRC budget adjustment. Sidney van den Bergh last month resigned as director of the Dominion Astrophysical Observatory in Victoria when the latest round of cuts was announced. He says the Victoria observatory has shut down one of its three telescopes, mothballed its optical shop and cut its support personnel to the absolute minimum.

Also struggling is TRIUMF (the Tri-University Meson Facility) in Vancouver, British Columbia. Last spring, NRC informed TRIUMF that its annual operating budget of \$28.1 million would be cut by \$4 million. Half that money was restored when it became apparent that TRIUMF would default on international contractual obligations without additional funds, but a \$24 million budget is likely for the next fiscal year.

Earlier this year, the three Canadian

Accord renewed

Washington

AFTER a five-year hiatus, Canada and the Soviet Union have signed an agreement to resume mutual academic, scientific and cultural exchanges. Canada suspended official exchanges in 1981 following the Soviet intrusion into Afghanistan. The new agreement takes effect at the start of next year, and will run until the end of 1988.

The Canadian move comes on the heels of a similar exchange agreement signed this past summer by the United States and the Soviet Union (see *Nature* 322, 491; 1986). Anton Yurkowich of the Ministry of External Affairs says Canada was the last member of the North Atlantic Treaty Organization to renew basic exchange agreements with the Soviet Union among those that suspended them after the arrival of Soviet troops in Afghanistan.

Yurkowich says the decision to renew formal contacts is not tied to any specific political change, but rather that "the time was ripe to get back to talk again". The Ministry of External Affairs has obtained budget authority for up to 27 individuals for academic and scientific exchanges, and is negotiating with the Association of Universities and Colleges of Canada to administer the programme.

The major beneficiaries will probably be Canadian Slavonic scholars who have been denied access to Soviet library and archive material. Assured access to archival material was also keenly sought by the United States in its accord. Joseph Palca

granting councils announced a unique scheme to encourage industry contributions to scientific research by matching private funds with government money (see *Nature* 320,102; 1986). Industry's willingness to participate in the scheme has yet to be demonstrated, however.

The Canadian Minister of State for Science and Technology, Frank Oberle, has promised to produce a national science and technology policy. With no firm timetable for its completion, many policy changes may be in place before the minister's plan is presented to parliament.

For those in government science posts both in and out of Ottawa, morale has plummeted. Many fear the current cuts are merely forerunners of larger cuts in the years to come, as the government shifts responsibility for basic research to universities and industry.

The furore created by NRC's cutbacks has prompted Oberle to form an independent panel that spokesman Dave Dale says will review the "whole overall macro picture" of NRC's role in Canadian science. Trying to sound optimistic, one NRC official said that although the council may be "more than slightly hurt, it is not mortally wounded".

Joseph Palca

Hubble Telescope taking shape

TECHNOLOGISTS from British Aerospace have just returned from California where they have been helping engineers from the European Space Agency's Technology Centre (ESTEC) to install the solar panels for the Hubble Space Telescope.

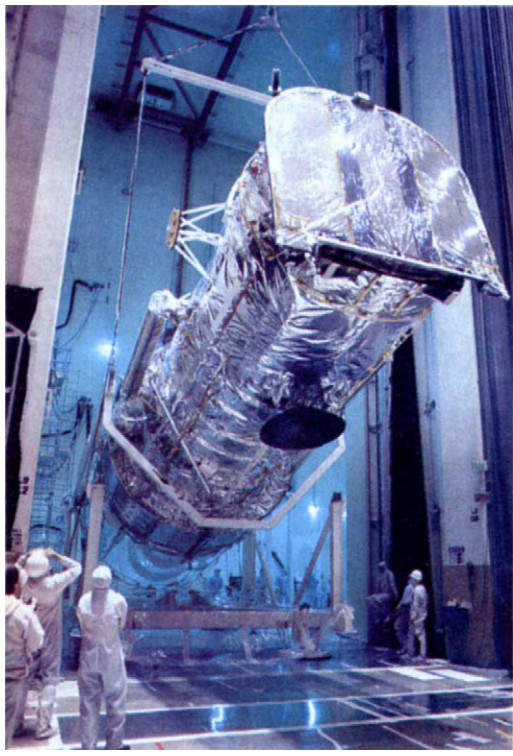
The arrays, built by a consortium led by British Aerospace, had been delivered in May to the telescope's construction site at the Lockheed Missile and Space Company complex in Sunnyvale. The arrays, the first to be designed to allow astronauts easily to retract and dismantle them, are contained on two panels unrolled from a single drum composed of more than 48,000 solar cells covering 356 square feet.

The retractability of the arrays and their ease of manipulation are two fundamental features of the design, as the telescope will have a 15-year life expectancy, three times that of the panels.

The telescope, due to be launched in two years time, will orbit at 593 km, allowing it to make observations free from Earth's atmosphere. It is designed to look seven times deeper into space, detect objects 50 times fainter and view them with 10 times greater clarity than the best telescopes on Earth. The Hubble Telescope will carry five scientific instruments that are replaceable and serviceable in orbit. They are the Faint Object Camera (FOC), the Faint Object Spectrograph, the High Resolution Spectrograph, the High Speed Photometer and the Wide Field/Planetary Camera.

Bill Johnstone

Hubble on test at Sunnyvale



Coral reefs

Okinawa runway under attack

Tokyo

THE battle to save one of the few thriving coral reefs left in Japan is about to reach its climax. The Okinawa prefectural government is expected soon to file a formal request with Japan's Construction Ministry to begin building an airport for jumbo jets on a reef of rare Pacific blue coral in Ishigaki Island, one of the southernmost islands of the Okinawa chain. Lined up to oppose the airport is an army of environmentalists from Japan and elsewhere.

The twin scourges of development and crown of thorns starfish (*Acanthaster planci*) have wreaked havoc with the coral reefs of the Okinawan Islands and the reefs of Ishigaki are rare remnants of what was once a thriving reef system.

Although Ishigaki already has an airport with a regular jet service to Okinawa, local government officials ten years ago proposed building a second airport with a longer 2.5-km runway to bring large jet-loads of tourists direct from Osaka and Tokyo, and in 1979 the site of Shiraho lagoon was selected.

Shiraho was selected because a nearby hill (Karadake peak) could be razed and dumped in the lagoon to create a level area of 130 hectares, conveniently walled in by the outer reef. But the people of Shiraho town (population 2,000) are vehemently opposed to the plan.

Two years ago, the Environment Agency planned a three-day survey of the reef. Local protesters tried to obstruct agency officials but were dissuaded by riot police and helicopters. Fearing further bad publicity, Ishigaki Town engaged Dr J. E. N. Vernon of the Australian Institute of Marine Science and Dr D. Kuhlmann of Humboldt University in Berlin to conduct a survey. But they were obstructed by Shoei Shirai, a local man appointed to help them, who later issued a report concluding that the reefs of north-west Ishigaki were the most valuable and the proposed site was not worth preserving. Vernon promptly disassociated himself from the report.

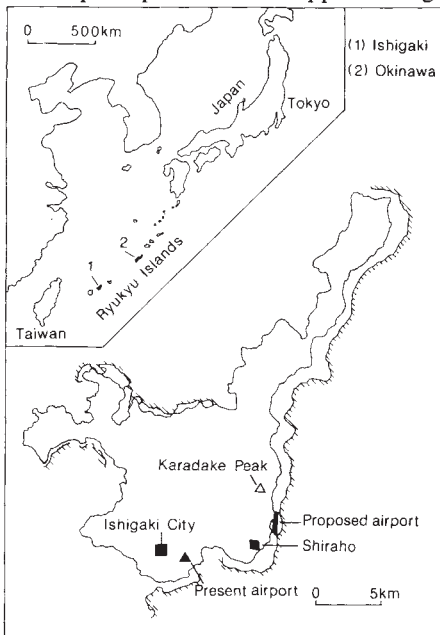
Vernon and other coral experts have been particularly struck by Shiraho's unusually extensive stands of Pacific blue coral (*Helipora cerulea*), a species listed under the Convention on International Trade in Endangered Species. Colonies tens of metres in diameter and 2-3 metres high stretch for hundreds of metres, constituting by far the biggest stands of this coral known in Japan.

In an attempt to allay concern, an informal advisory panel to the Governor of Okinawa that recommended construction of the airport has suggested that the coral be removed and 'replanted' elsewhere. But coral is not the only concern of oppo-

nents to the airport. Many have questioned the necessity of a second airport for such a small island, and there are fears that the 'spare' airport may be used for military purposes—Ishigaki is strategically placed at the southernmost tip of Japan.

Opposition to the airport is attracting worldwide attention. Last summer Jacques Cousteau wrote to the governor of Okinawa appealing for the plan to be dropped, and last month more than 500 environmentalists from Japan, the United States and the Philippines chartered a ship and called at Ishigaki to hold a protest rally on Shiraho beach.

In a ploy to delay construction, several hundred square metres of land needed for the airport apron that had slipped through



the government buyer's hands were promptly bought up by the anti-airport group and sold off in doormat-sized lots to supporters. The government will now have to negotiate with about a thousand 'owners' to get the land back.

But the biggest obstacle for the airport planners is the fishing industry of Shiraho. Fishermen in Japan wield extraordinary power, and their consent must be won before construction projects affecting fishing grounds can go ahead.

Although the 600-member fishermen's cooperative of Ishigaki accepted ¥500 million (£2 million) in 1980 as compensation for loss of fishing grounds at Shiraho, of which ¥200 million was earmarked to pay off the cooperative's debts, the 30 fishermen of Shiraho have steadfastly refused to accept the agreement, and have taken their cooperative to court. As long as the disbursement of compensation remains blocked, airport construction cannot go ahead.

David Swinbanks

Yugoslavia

Serbian centenary "postponed"

SERBIA'S Academy of Sciences and Arts, embroiled in a fierce political row with the authorities, has postponed the celebration of its centenary, due next week, until May. This decision follows fierce criticism by party and government leaders of an outspoken memorandum on the continuing social, economic and political crisis in Yugoslavia prepared by an academy group headed by Dr Antonije Isakovic, a vice-president of the academy. Under pressure, the academy has stopped work on the project.

Both the memorandum and the academy's decision to postpone its celebrations have been criticized. The Belgrade daily *Politika* on 23 October described the decision to postpone the celebrations as a kind of blackmail, and said that neither a handful of academics nor the president of the academy has the right to alter the date of the celebrations, which was "fixed by the history of its existence". *Politika* seems to have forgotten that, in 1974, the Soviet Academy of Sciences found it necessary to postpone the celebration of its 250th anniversary for a year.

The disputed memorandum deals with all aspects of Yugoslav life and presents a grim picture of bureaucracy, arbitrary decision-making, corruption, nepotism and a disregard for the law by those in positions of authority. Its advocacy of constitutional reform and the genuine participation of the public at large in decision-making has been attacked by the party

theorists as anti-socialist, a betrayal of the legacy of Tito and evidence of a wish to break up the Yugoslav federation.

On education, though, the memorandum seems more concerned with keeping Yugoslavia together than with dismembering it. It castigates as divisive the present set-up in which each of the six republics and two autonomous provinces that constitute the federation has its own educational system (and academy of sciences). The memorandum says that countries without unified educational systems cannot expect to remain unified.

The memorandum also attacks a number of the educational concepts beloved of the party theorists, including that of "job-oriented" education, the result of which has been the over-production of ill-trained "specialists" unable to play a part in the social and economic life of Yugoslavia. It also says that the party campaign against elitism has meant a levelling down of standards and that the requirement that all persons employed in education should be "morally and politically suitable" has made the universities hot-beds of "careerism and conformity".

The Isakovic group also attacks the much-praised reform of the universities which hived off research into separate institutes, saying that the result has been the destruction of the laboratory base of the universities and the compartmentalization of science. The memorandum concludes that the universities therefore cannot play

their essential part in the "technology-computer revolution". It advocates a thorough reorganization of scientific institutions in the interests of modernization, efficiency, greater financial investment and freer creativity.

Vera Rich

X-ray astronomy

British-Soviet collaboration

A JOINT British-Soviet X-ray astronomy satellite may be launched in the early 1990s. Five British X-ray astronomy groups are now working out proposals for experiments to be submitted to the Soviet side at the end of December. If all goes well, the Soviet acceptance is likely to come by the middle of next year. The project, generated by a visit of a delegation from the British National Space Centre to the Soviet Institute of Space Research at the beginning of October, is viewed by both sides as an important development. Dr Roald Sagdeev, director of the institute, said the meeting showed that Soviet and British scientists have good prospects for cooperation in space, especially in astrophysics, radioastronomy and the study of materials in space.

Until now, there has been relatively little direct cooperation between the two countries in space. The most significant project, so far, has been the British involvement in Roentgen, a module of four large X-ray telescopes to be mounted on the Soviet Mir orbital station next year.

But it was not the Soviet Union that initially proposed that Britain should take part but the Dutch, who wished to work in the hard X-ray/low-energy gamma-ray range using a coded-mask telescope, the principle of which was pioneered at the University of Birmingham. The first such telescope was flown on the Skylark rocket in 1976, and an improved version aboard Spacelab-2 in 1985. Professor Peter Willmore from Birmingham, who is now coordinating the proposals for the satellite, is reluctant to say too much about possible experiments. But the final proposal, he said, would be for a cluster of telescopes, using a variety of techniques to cover various energy bands.

The satellite, if all goes well, will be financed by each country paying for its own experiments and the Soviet Union covering the cost of the satellite and the launch. The project would therefore follow the normal financial pattern of Soviet collaborative projects, and has nothing in common with the purely commercial proposals, made by the Soviet Union in the wake of the Challenger disaster, that the Soviet Proton rocket could be used to launch satellites for foreign customers.

Vera Rich

Ups and downs of French science budget

THE French research budget for 1987, up on the budget for 1986 as corrected after the present government came to power in April, but down compared with that planned for 1986 by the previous government, seemed to cause more excitement in the French National Assembly last week than any other aspect of next year's planned spending.

Leading socialist Jean-Pierre Chevènement was reportedly in fine form. His grand edifice of a ministry for research, technology and industry has been gradually dismantled, first by his own government (which trimmed off industry) and then by the present incumbents (technology has now gone, and the rump has been attached to the ministry of education). The ex-minister was exacting his revenge. He told the present science minister, Alain Devaquet, who is proud of his achievement in protecting the research councils and winning a modest rise in budgets against considerable internal opposition, that his budget was "faked and inflated". It was a "conjuring trick ... unworthy" of a sym-

pathetic man like Devaquet.

The new minister rose to defend himself. The use of a single quantity, finance, was not the way to handle such a complex matter as French science, he said. Nevertheless, the national budget in government and industrial research and development would be up 8.5 per cent (on actual 1986 spending) to FF127,000 million (around £13,000 million), Devaquet claimed. A particularly big increase in defence research spending in the 1987 budget would offset decreases in support for industrial research in other areas. Criticized for this by socialists, Devaquet stoutly defended defence research: "I don't share the disdain of certain speakers for military research . . . It has done a great deal for many important industries."

However, André Giraud, the defence minister, will now have a research budget of FF30.8 million. Devaquet's budget for basic science is FF21 million. "So who is now the minister of research?" Chevènement taunted. Well, not M. Chevènement, anyhow . . .

Robert Walgate

Semiconductor rivalry

US bid to equal Japanese

Washington

THE beleaguered US semiconductor industry, hard pressed by superior manufacturing technologies from Japan, hopes to regain its competitive edge in world trade through a major cooperative manufacturing enterprise called Sematech. The Semiconductor Industry Association (SIA) will meet next week in Long Beach, California to give formal approval to the "manufacturing excellence project" that is already getting strong encouragement from US government.

Warren Davis of SIA says the industry's plight stems from the inability to develop effective manufacturing technologies for implementing innovative semiconductor designs. Davis believes no single US company is capable of taking on the scale of project needed for creating a new generation of manufacturing technologies.

Since early summer, representatives of the semiconductor industry have been holding discussions among themselves and with government officials on how best to proceed with Sematech. Some companies would like to see a full-scale manufacturing project, with the manufacturing muscle to beat the Japanese. Others would like a smaller pilot programme, with the technical fruits falling to individual companies. A compromise between the two approaches is likely.

Concern in Washington has been growing about the health of the US semiconductor industry, with several recent reports raising national security alarms over availability of semiconductor products (see *Nature* 323, 483; 1986). Davis says if Sematech is to work, the government will have to give assurances that the effort will not be hindered by antitrust prosecution or unfavourable tax status. Referring to the Sematech project, Bruce Merrifield, assistant secretary of commerce for productivity, technology and innovation has spoken enthusiastically about the government's role as "the catalytic element of change", eliminating potential barriers to success. SIA officials were in Washington last week for talks with Merrifield and National Science Foundation (NSF) director Erich Bloch concerning Sematech. Although no government funds have been committed to the project as yet, commitments on taxes and antitrust issues could be worth millions of dollars. The Semiconductor Research Corporation, an industry supported consortium, recently received a grant worth \$1.2 million from NSF and the Department of Defense, and could be a conduit for additional government funds.

Joseph Palca

Ozone depletion

US gets tough on CFC emissions

Washington

THE US government startled environmental conservatives and delighted radicals last week by calling for a complete halt to global emissions of halogenated alkanes thought to be destroying stratospheric ozone. The proposal was made in a paper on principles for an international protocol for protection of the stratospheric ozone layer issued by the State Department.

The United States will take the proposal as a starting position for the international negotiations beginning at Geneva on 1 December intended to put flesh on the bones of the Vienna Convention on Protection of the Ozone Layer. The convention has been signed by 25 countries since March 1985, but negotiations on a protocol to implement controls on chloro-

fluorocarbon (CFC) emissions have been postponed until the completion of informal technical meetings.

Anxiety about damage to the ozone layer, which filters out high-frequency ultraviolet radiation from the Sun, has grown in recent months. A 'hole' in the ozone layer appearing over Antarctica during the Antarctic winter (see *Nature* 321, 729; 1986 and 315, 207; 1985) is just now the subject of a study mounted by an inter-agency US team.

Although the cause of the 'hole' is not definitely known, the team's preliminary observations suggest that "a chemical mechanism is fundamentally responsible for the formation of the hole". Global ozone cover is thought to have decreased by 2-3 per cent in 1970-80, although measurements have been hard to interpret. CFCs also contribute to the greenhouse effect, by which the lower atmosphere may be warmed because of the effects of pollutants on the atmospheric radiation balance.

According to the State Department's proposal, a "long-term scheduled phase-out" of emissions should follow a near-term freeze on emissions of all fully halogenated alkanes. These include the widely used chlorofluorocarbons CFC 11, 12 and 113, believed to be the worst offenders, and halons 1211 and 1301. The proposal says that present uncertainties preclude the definition of a safe level of emissions, but that given the risks to health and the environment, the measures are justified. The proposed progressive reductions would be subject to change in the light of accumulating data.

The US proposal mentions in passing the need to develop incentives to persuade non-signatories of the convention to participate in the protocol.

The US industry group of chlorofluorocarbon manufacturers, the Alliance for Responsible CFC Policy, claims that the US proposal is not based on scientific principles. The alliance has proposed a cap on CFC production capacity and voluntary emission reduction measures to meet concern about the ozone layer.

The US proposal was based at least in part on a draft study by the Environmental Protection Agency, which estimates that incidence of skin cancer could increase significantly if emissions continue unchecked. But the industry alliance points out that the study has not yet been formally reviewed by the agency's scientific advisory board. A spokesman said that a phasing out of emissions did not necessarily mean an end to production, if some means could be devised to return CFCs to manufacturing plants for reuse.

Tim Beardsley

Deaths foretold

Washington

AN extra 40 million skin cancer cases and 800,000 deaths among people now alive and those born before 2074 are likely in the United States if CFCs are not further controlled, according to EPA's draft study* of stratospheric modification. The predictions stem from the expected increase in ultraviolet-B radiation reaching the Earth's surface as stratospheric ozone declines. The EPA study is largely responsible for the tough new US position on emission controls.

The predictions assume annual growth of 2.5 per cent for CFCs and 0.6 per cent for carbon dioxide, in line with current estimates. They also assume a continuation of present growth rates of nitrous oxide and methane (0.25 per cent and 1 per cent respectively). Equilibrium global temperature would under these conditions increase by between 3 and 9.5 degrees celsius by 2075, assuming the National Academy of Sciences' estimate of a 3 degree rise from a doubling of carbon dioxide.

EPA says that, as well as increasing skin cancers and cataracts, the increase of ultraviolet-B could promote the spread of herpes and leishmaniasis. There may also be effects on aquatic food chains and terrestrial plants, although information is scanty. But the study shows that estimates of risks are highly sensitive to CFC and halon growth rates; if the CFC growth rate is halved from 2.5 per cent to 1.2 per cent, estimated damage is reduced by 90 per cent. If a growth rate of 3.8 per cent is assumed, estimated damages increase by 400 per cent.

Tim Beardsley

An Assessment of the Risks of Stratospheric Modification (final draft), submission to EPA Science Advisory Board, October 1986.

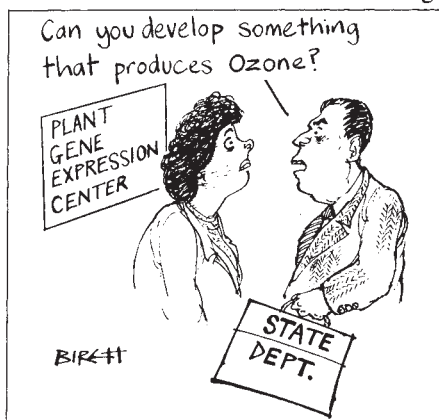
US agricultural research

Plant genes now fashionable

Washington

THE US Department of Agriculture (USDA), frequently criticized for its neglect of plant molecular biology, is now preparing to open a showcase institute with this theme at Albany, near San Francisco, California. The new Plant Gene Expression Center could herald a growth of interest from other government agencies.

At the outset, the new centre, under the aegis of USDA's Agricultural Research Service (ARS), will work on the transformation and regeneration of monocotyledons, the use of transposons to 'tag' and identify genes, molecular mechanisms of hormone action and environmental reg-



ulation of gene expression. USDA's annual support will be about \$6 million, but extra funds will be sought from other sources.

The scientific director will be Peter Quail, now at the University of Wisconsin. The centre will eventually house 10 research groups, some of whose leaders will also be professors at the nearby Berkeley campus of the University of Califor-

nia. Quail says the Berkeley link will be important; staff will be employees of the university, not civil servants, and visiting scientists will play a vital role. He hopes the centre will develop techniques for manipulating plant genes coding for informational molecules that will have long-term applications in crop improvement.

But while USDA is planning its own modest expansion of basic plant biology, the White House's Office of Science and Technology Policy is pushing for a more far-reaching increase in federal support. The view there (shared elsewhere) is that plant science has been eclipsed by the popularity of biomedicine, and urgent action is needed to redress the balance.

Indeed, the Office of Management and Budget hopes that there will be funds in the 1988 federal budget for an inter-agency initiative in basic plant biology. This proposal has been made by the Departments of Agriculture and Energy and the National Science Foundation and would support several new 'interdisciplinary research groups' in plant biology at universities, concentrating in five areas:

- microbial ecology and genetics in the rhizosphere
- chemistry of complex plant carbohydrates
- molecular biology and genetics of plant cells
- modelling of agro-ecosystems
- information technology.

Part of the impetus for the new initiative is the need to develop a 'predictive ecology' to help resolve the regulatory paralysis over environmental releases of genetically modified organisms. Whether those who hold the purse strings will support a major expansion remains to be seen.

Some academics are sceptical about the plan, saying that researchers can collaborate when they need to as things are and that the plan may simply be a way of dressing up a case for more money. Virginia Walbot of Stanford University believes



Peter Quail: aiming at better crops despite agricultural surpluses.

that universities will find it hard to assemble true interdisciplinary groups without the promise of long-term support, which may be hard to get.

Others point out that the director of the Office of Science and Technology Policy, William Graham, whose support would be vital, has a background in defence, and may not be enthusiastic about biological research.

On a political level, the problems of US farmers are seen as stemming more from surpluses than from shortages, and biological research is seen — rightly or wrongly — principally as a way to make the surpluses bigger. **Tim Beardsley**

FRG-Soviet agreement

German deal put on ice

Hamburg

THE proposed agreement on research collaboration between West Germany and the Soviet Union seems to have been one of the first victims of the ill-feeling generated by Chancellor Helmut Kohl's controversial reference last month to Mr Mikhail Gorbachev, the general secretary of the Soviet Communist Party.

In an interview with the US news magazine *Newsweek*, Kohl had described Gorbachev as an "expert in public relations" and had likened him to Hitler's minister of propaganda, Josef Goebbels. Now Moscow has cancelled a visit during which Dr Heinz Riesenhuber, minister of research and technology in the federal government, was to have signed an agreement on the peaceful uses of nuclear energy and other matters reached only in July this year (*Nature* 322, 488; 1986) after a decade of intermittent negotiation.

As things are, there is an exchange agreement between the Deutsche Forschung Gemeinschaft (DFG) and the Soviet Academy of Sciences but no agreement on scientific collaboration between the West German and Soviet governments.

Jürgen Neffe

Grassroots unrest in plant biology

Washington

DESPITE the creation of the new Plant Gene Expression Center, seen by many as a belated attempt by (USDA) to jump on the plant biotechnology bandwagon, many academics believe research in basic plant biology has been seriously neglected.

The figures are telling; ARS's total budget will be \$498 million this year, but that is all in-house and includes the cost of running laboratories. At least one-half is applied research such as plant breeding.

The Cooperative State Research Service, the USDA research arm which, jointly with state governments, supports the state experimental research stations, has a budget of \$300 million, but only \$40 million goes as competitive grants open to the university community. On one estimate, total federal support for innovative basic

plant science open to competitive proposals from academics, including programmes of the Departments of Agriculture and Energy and the National Science Foundation, is only \$110 million, compared with over \$3,000 million spent on basic research by the National Institutes of Health in 1986.

ARS does support many small projects that use modern molecular biology techniques, but critics, such as Roger Salquis, president of Calgene Inc., and Stanford's Virginia Walbot agree it could have made better use of its opportunities. And the minuscule competitive grants programme, though increased in size in recent years, is not popular with academics; fewer than 10 per cent of proposals succeed, while grants (including overhead) vary between \$50,000 and \$80,000 in amount and last for only two years.

Tim Beardsley

Remedy for British science

SIR—I agree with the spirit of your leading article (*Nature* 323, 655: 1986) bemoaning the plight of British science, but I would question the political utility of the 'precepts' you urge the Advisory Board for the Research Councils (ABRC) to adopt to promote its case. In many ways they are all too familiar claims that those defending science have used in its defence. As such, they are as likely to be as unsuccessful in persuading Whitehall that its 'policies' — if that is not to dignify the disjointed initiatives of recent years too much — are damaging science as has been the more straightforward claim that 'we need more money'.

In many ways your precepts understate the real damage that is being done to the infrastructure of British science research. ABRC would be better advised, I believe, to make more of the wider national economic context in which science research and development finds itself, as well as to draw on what detailed information there is about the real damage that is being done to science at 'grass-roots level', the fundamental knowledge-generating culture of science.

On the first point, ABRC would do well to point out to its political masters that, according to the government's own Central Statistical Office, British industry, excluding oil, grew by only 1.5 per cent in 1979–85: little wonder that industry has had to look 'innovatory gift-horses in the

mouth'. Moreover, where growth has occurred in the economy — in the financial sector — there appears to be little incentive on the part of financiers to invest in innovative research and development, basic, strategic or applied: money is best made out of money, on the financial markets.

At the more local level of laboratory life, my own research into the state of plant biotechnology and plant breeding in the United Kingdom reveals a failure by government to understand the research interests, abilities, needs and capacities for innovation of plant science. The speed and multidirectional nature of initiatives including privatization, rationalization and reorganization tend to pull (and literally so for its staff) science in opposite directions, breaking up the critical mass of research so vital to the future. Given the broader economic context sketched above, it is unlikely that scientific research and development is moving sideways into an expanding world of commercial opportunity. Management of science does need to be improved at national and local level: but clearly it can be effective only if it is informed by and sensitive to both.

A.J. WEBSTER

*Plant Biotechnology Project,
Cambridgeshire College of Arts and
Technology,
East Road,
Cambridge CB1 1PT, UK*

Pressure to publish

SIR—Writing in your New Journals Review issue, Mary Holmes admirably summed up the problems we face with the ever-increasing range of specialized scientific periodicals available (*Nature* 323, 359–360; 1986). A cause contributing to this information explosion is the extraordinary pressure on scientists to publish. This is certainly noticeable in the medical field, where appointments and promotions frequently hinge on the sheer number of papers published in reputable-sounding journals. (A letter to *Nature* might get more brownie points than a paper in *The British Journal of Controversial and Spurious Knowledge* but 70 papers are more likely to get you a senior lectureship than one seminal paper.) This puts pressure on scientists to subdivide pieces of research to produce a number of papers rather than one complete publication.

The recent attempts in Britain to assess the research worth of entire departments or institutions on the strength of five publications has not proved particularly happy, but I do not see why the abilities of an individual should not be assessed on the strength of half-a-dozen papers. If

enough institutions changed that part on the back of their job application forms which says "Publications" to "List not more than six of your publications" the pressure on scientists would change to trying to produce a small number of very good publications. If this happened, a lot of semi-vanity publications would cease, and much material currently appearing in glossy print would be reduced to a footnote saying something like "Further details of this experiment are obtainable from the author on request" with no great loss to the flow of scientific information.

MARTIN GUHA

*Institute of Psychiatry Library,
De Crespigny Park, Denmark Hill,
London SE5 8AF, UK*

Scots elders

SIR—In the Opinion article "British science over the hill" (*Nature* 323, 655; 1986), you say that three Scottish universities are more than 150 years old. The true figure is four: St Andrews (founded 1411), Glasgow (1451), Aberdeen (1495) and Edinburgh (1583).

A.M. HETHERINGTON

*113 Scotforth Road,
Lancaster LA1 4SD, UK*

Isolated South Africa

SIR—*Nature* has a sound record (see 322, 115; 1986) on human rights and freedom from repression for academics, but the conditions in South Africa would probably not have persisted if the grandees of the scientific community and the other learned academic communities had deployed their corporate influence against apartheid. In general, scientists and academics enjoy great prestige and privilege, and exert considerable influence within their own institutions. A large part of scientific research is supported by the public purse, either directly from government grants or from charitable foundations. It is not difficult for scientists to retreat into their laboratories until their research grants are up for renewal.

The cause of international scholarship is not more important than the fate and dignity of some 23 million black South Africans and five million white South Africans. One characteristic feature of the South African regime is its sensitivity to isolation. So long as South African athletes, diplomats, scholars and other professional and business people are granted free access to countries outside South Africa, especially the United Kingdom, Canada and the United States, the present regime in Pretoria knows that it will rule the Cape for a very long time.

The arguments against sanctions are based on previous experience when there was no common consensus for sanctions among the major European governments, Canada and the United States. There now seems to be a growing trend towards such a consensus.

J.A. BARNES

*Department of Oral Biology,
University of Alberta, Edmonton,
Alberta, Canada T6G 2N8*

Titles of references

SIR—I agree with John Maddox (*Nature* 322, 681; 1986) that scientific literature is poorly written for the communication of the essential ideas of the research carried out.

One point he neglected to note is that a major contribution to this state is the obscure documentation of references. It is my opinion that a complete reference should always include the title of any cited work. I find this omission a great hindrance when researching related topics because I cannot determine beforehand whether an article is even remotely connected to my needs.

Unfortunately, I note that *Nature's* reference style omits the titles of cited articles from journals. Your editors are contributing to this lack of communication as much as the authors.

BRIAN G. LATHAM

*General Delivery, Yellowknife,
NWT, Canada X1A 2L8*

Cytoskeleton

Dynamic microtubule dynamics

Robley C. Williams Jr, Michael Caplow and J. Richard McIntosh

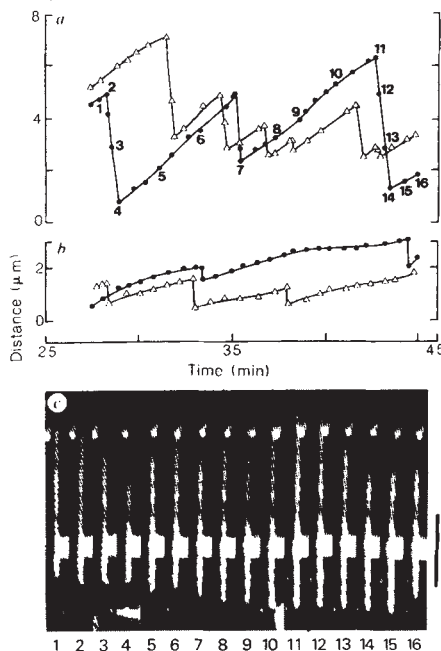
WHEN assembly and disassembly of the microtubules from eukaryotic cells was first accomplished *in vitro* thirteen years ago, most observers thought the processes involved a simple reversible equilibrium between the polymeric microtubule and its tubulin subunits. Recent investigations reveal a livelier side of microtubule behaviour: an active growth and shrinkage of the polymer. A vivid example is the work of Horio and Hotani¹, who observed untreated microtubules *in vitro* in real time and found striking fluctuations in their lengths.

Microtubules in living cells come in all flavours of stability, from the ephemeral polymers of the mitotic asters to the sturdy stuff of which cilia and the protozoan cortex are constructed. Regulation of this diversity is, as yet, poorly understood, but it seems certain that the modes of microtubule assembly and disassembly lie at the heart of the matter. The assembly of tubulin subunits is accompanied by the hydrolysis of GTP. This reaction can produce a steady-state flux of subunits through the microtubule, a process termed treadmilling (see the discussion in *News and Views*²). Hydrolysis of bound GTP can, however, lag behind the incorporation of subunits³. The model that follows from this observation predicts the formation of inhomogeneous microtubules. In microtubules that are either growing or at a steady state of polymerization, the core of the microtubule is expected to consist almost completely of GDP-tubulin, whereas the recently added tubulin at the ends should form a GTP-tubulin cap^{3,4}. Dissociation of a GDP-tubulin from the end of the exposed core is thought to be about 600 times faster than that of a GTP-bearing subunit from the cap⁵. Consequently, a microtubule is metastable: rapid disassembly results if the GTP-cap is destroyed (for example by dissociation of subunits from the end, by breakage, or by GTP-hydrolysis propagating endwards). According to this model, the macroscopic steady state of polymerization results from occasional rapid shrinkage of some microtubules, balanced by the slow growth of others.

This concept of microtubule behaviour was graphically supported by observations on the assembly of pure tubulin onto naturally occurring microtubule organizing centres⁶. Mitchison and Kirschner found that some of the polymers can shrink to disappearance while their immediate neighbours are growing. This phenomenon was presumed to occur via removal of subunits from, and addition to,

the ends of the microtubule. Microtubules grown by elongation of preformed fragments also displayed a striking, simultaneous decrease in number and increase in the mean length of those remaining, both at steady state and on dilution. In the 'dynamic' instability model put forward to explain these observations, a given microtubule is in either a growing or a shrinking condition (phase), but is probably never static⁷. The relative rates of microtubule growth and GTP hydrolysis are thought to be the most significant factors in maintaining the GTP-cap⁷. (See ref. 2 for a discussion in *News and Views*.)

Is this model thermodynamically and kinetically sound? Is there perhaps some simpler alternative? The first question has



Interconversions between the growing and shortening phases at the ends of a single microtubule. *a*, Length change at active end; *b*, change at inactive end; *c*, successive frame of a video image from one microtubule. (From ref. 1.)

been answered positively⁸ in model studies that cast assembly and disassembly in terms of elementary reactions between soluble tubulin and plausible representations of the microtubule lattice. By choosing hypothetical values for the rate-constants, the model mimics the dynamic behaviour of microtubules in solution. The second question will not be as easily answered. Mitchison and Kirschner^{6,7} explained the simultaneous increase in length and decrease in number of microtubules by an endwise, total disassembly of some polymers and addition of the released subunits to other, growing, micro-

tubules. But immunogold labelling and electron microscopy studies⁹ show that segments of preformed microtubules can also join together end-to-end to form a longer microtubule. This seemingly improbable but evidently quite efficient reaction, observed also with uncapped, taxol-stabilized, GTP-free microtubules (R.C.W. and M.C., unpublished observations), may contribute to the length redistributions on which the dynamic instability model is based.

Any doubt, however, that microtubules at steady state do fluctuate in length has now been dispelled by Horio and Hotani¹. By means of real-time observation with dark-field optics, they can watch single microtubules change length (see figure). A given microtubule, when it is attached to the surface of a glass coverslip, can be followed through many cycles of shrinkage and growth. Their observations show directly that either end of each microtubule can exist in a growing or shrinking phase but that the length of a microtubule is never stationary. Micrometres of polymer, corresponding to thousands of subunits and perhaps 10–90 per cent of the length of a microtubule, assemble or disassemble between phase changes. The processes occurring at the two ends of one microtubule behave independently.

Not all these new results provide simple support for the picture developed by Mitchison and Kirschner^{7,8}, who explained the decrease in microtubule number and increase in microtubule length by asserting that the depolymerization phase would generally continue to completion, causing a microtubule to disappear from the population. Horio and Hotani's direct observation reveals no disappearing microtubules. Phase-reversal from shrinking to growth must be sufficiently probable, under the conditions of observation, to rescue essentially all microtubules from vanishing. One might think that this lack of conformity could be caused by the chamber used to make the observations. Attachment of microtubules to the glass surface might block their complete disassembly. But if this were the case, other microtubules not so close to the glass would disappear, thereby increasing the concentration of tubulin above that available at steady-state, so that the microtubules seen in the microscope would continually elongate. Because they neither disappear nor elongate, something other than simple dynamic instability must be going on.

Horio and Hotani also find, unexpectedly, that the end which undergoes more frequent reversals of phase, the 'active' end, grows and shrinks more rapidly than the infrequently reversing 'inactive' end. In current models, one imagines a GDP-bearing central region of the microtubule that gradually becomes longer as GTP hydrolysis proceeds on the incorpo-

rated subunits. At the active end, addition of GTP-bearing subunits should be more likely to outrun the spreading hydrolysis, leading to less frequent phase reversal than at the inactive end. It may be possible to reconcile the observed results with the GTP-cap model if GTP-hydrolysis occurs predominantly at the interface between the GTP-tubulin cap and the GDP-tubulin core (as suggested in ref. 10) and most rapidly at the active end.

Microtubule-associated proteins bind to the outside surface of microtubules, altering the energetics and kinetics of their assembly and disassembly. It seems likely that the GTP-cap persists in the presence of microtubule-associated proteins⁵. Yet when Horio and Hotani add these proteins to their microtubules, the frequency of phase reversal is greatly reduced. This result appears to be related to the observation¹¹ that microtubules containing microtubule-associated proteins do not show the length redistributions seen by Mitchison and Kirschner^{6,7}, but do undergo the treadmilling process. The function of microtubule-associated proteins in modulating the dynamic instability of

microtubules thus may be physiologically important. This mechanism may, for instance, underlie the large, cell-cycle-dependent differences in lability of microtubules observed in photobleaching experiments with tubulin¹². □

1. Horio, T. & Hotani, H. *Nature* **321**, 605 (1986).
2. McIntosh, J.R. *Nature* **312**, 196 (1984).
3. Carlier, M.-F. & Pantaloni, D. *Biochemistry* **20**, 1918 (1981).
4. Hill, T.L. & Kirschner, M.W. *Int. Rev. Cytol.* **78**, 1 (1982).
5. Carlier, M.-F., Hill, T.L. & Chen, Y.-D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 771-775 (1984).
6. Mitchison, T. & Kirschner, M. *Nature* **312**, 232 (1984).
7. Mitchison, T. & Kirschner, M. *Nature* **312**, 237 (1984).
8. Chen, Y.-D. & Hill, T.L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4127 (1985).
9. Rothwell, S.W., Grasser, W.A. & Murphy, D.B. *J. Cell Biol.* **102**, 619 (1986).
10. Caplow, M. & Reid, R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3267 (1985).
11. Farrell, K.W., Jordan, M.A., Miller, H.P. & Wilson, L. *J. Cell Biol.* (in the press).
12. Saxton, W.M. *et al. J. Cell Biol.* **99** 2715 (1984).

Robley C. Williams Jr is in the Department of Molecular Biology, Vanderbilt University, Nashville, Tennessee 37235 USA; Michael Caplow is in the Department of Biochemistry, University of North Carolina, Chapel Hill, North Carolina 27514, USA; J Richard McIntosh is in the Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, Colorado 80309, USA.

Natural selection

Rapid evolution of urban birds

Jared M. Diamond

A MAJOR dispute in evolutionary biology concerns the rate at which species evolve to occupy new niches. How suddenly did *Homo sapiens* replace *H. erectus*, and how soon after the demise of dinosaurs did large mammals arise? It is inevitably difficult to answer such questions using fossil evidence. Yet we live among many little-studied cases of rapid evolution, as species occupy habitats created by man. An example is the very recent formation of an urban avifauna in Papua New Guinea, summarized in a posthumously published paper by Harry Bell (*Proc. Western Fdn Vert. Zool.* **3**, 1; 1986).

In Europe, where towns have existed for thousands of years, many species of birds have become human commensals. Rock doves, starlings and house sparrows thrive in concrete habitats, and the latter species is now so closely associated with man that its original niche is unknown. In Papua New Guinea, however, the first town was founded a mere 112 years ago, and today only 16 towns have human populations exceeding 4,000. Typical town habitats are very unlike the New Guinea rainforest and consist of scattered trees and ornamental shrubs, mostly of introduced species. How quickly are bird species adapting to these new habitats, and what makes for a successful colonist?

In some cases the pre-adaptations are clear. The greater wood-swallow, which

traditionally sallies from tall ridge-top trees over the primary rainforest, in towns sallies from radio towers and telephone poles. Shining starlings transferred their nests from holes in trees to vents of air conditioners, thereby becoming a pest. Swiftlets that would otherwise nest in caves now use tunnels excavated by the Japanese army during the Second World

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Frank Lane.

Urban rock doves (feral pigeons).

War. A nocturnal ground-feeding predator, the Papuan frogmouth, has learned to hunt rats under street-lights.

Such shifts in behaviour do not occur instantly and newly acquired behaviours take time to spread. In several cases the timetable is known. It was not until 1971-74 that a flock of black kites settled in the town of Lae to feed on road kills, although the town was founded in the 1930s. Only since 1983 have brahmyn kites learned to feed on road-killed toads

in Port Moresby. Goldie's lorikeet, which was formerly a rare species occurring in primary forest, arrived in the 1970s to feed on casuarina seeds in highland towns, where it is now the most abundant bird. Lemon-bellied flycatchers expanded from the savanna into the town of Wau between 1976 and 1978, 50 years after Wau was founded.

Checkered distributional patterns resulting from interspecific competition are common among closely related New Guinea bird species (my own work, in *Ecology and Evolution of Communities* eds Cody, M. & Diamond, J.M. p. 342; 1975). Bell's analysis of the town distribution of fruit doves (genus *Ptilinopus*) gives the first indication of how quickly the checkerboards develop. Three lowland species (orange-bellied, orange-fronted and pink-spotted) are each now resident in one or more towns, but no town except Popondetta supports more than one of the species, so that the distributions form a checkerboard. In natural habitats the orange-bellied and pink-spotted fruit doves are geographically widespread, the pink-spotted extends to higher elevations, but the orange-fronted is absent from the region that includes the towns of Lae and Madang. Town distributions are as follows: the orange-fronted is common in all lowland towns throughout its geographical range; the orange-bellied is found as a town bird only in Madang and Lae (and in Popondetta where it is less common than the orange-fronted and is possibly declining); and the pink-spotted occupies only Bulolo, above the altitudinal ceiling of the orange-fronted and orange-bellied. These results suggest that the orange-fronted has a competitive advantage in towns, that the orange-bellied succeeds only in the absence of the orange-fronted and that the pink-spotted succeeds only in the absence of both relatives. If so, this distribution must have arisen within the past 40 years and may not yet have reached equilibrium in Popondetta. There is a similar distribution in New Zealand, where two introduced starlings (the European starling and the Indian myna) have both acquired the habit of surveying pastures for food while perching on the back of a sheep, but in any given pasture only one of the two species occupies the sheep.

The successful town colonists are unrepresentative of the New Guinea avifauna as a whole in their preferences among natural habitats. Of species present in natural habitats near towns, only 4 per cent of the rainforest species but 33 per cent of the savanna species and 50 per cent of the second-growth species occupy towns. These differences partly reflect the varying similarities between the natural habitats and town habitats (least like rainforest), partly the behavioural plasticity of second-growth species. The

very low success of rainforest species has grim implications, because it means that spreading urbanization will threaten New Guinea's famous rainforest avifauna.

As Bell notes, New Guinea town avifaunas are still far from the equilibrium that now exists in Europe. Changes in behaviour and status are occurring rapidly. The species/abundance relation remains unusually equitable: many species are vying for space in New Guinea towns, in contrast to the pattern of a few dominant species in older towns in other countries. This panorama of biology in rapid flux provides opportunities for field biologists to address some basic questions. Does the sudden acquisition of new behaviour depend on a genetic change or on culturally transmitted learning? How quickly does

morphological and physiological evolution occur among colonists of the new habitats? We know that house sparrows living in wet or dry and hot or cold areas of North America differed morphologically when studied a century after their arrival (Johnston, R.F. & Selanders, R.K. *Science* **144**, 548; 1964) and that the advance of house sparrows in tropical Australia halted for 60 years before their final surge across the Tropic of Capricorn. Biologists neglected those and innumerable other opportunities to study natural selection in action. Such opportunities still exist in New Guinea and elsewhere. □

Jared M. Diamond is Professor of Physiology at the University of California Medical School, Los Angeles, California 90024, USA.

Metallic superlattices

The supermodulus syndrome

Robert W. Cahn

SYNTHETIC compositionally modulated epitaxial semiconductor structures, or semiconductor superlattices, were first proposed by Esaki and Tsu in 1969. Because of their remarkable electronic and optical properties, these structures have been investigated intensively since then. The metallic equivalent of semiconductor superlattices, synthetic modulated metal films, began to excite interest in 1977, but until now studies on these films have been scientific rather than technological. Recent work has succeeded in using synthetic, bimetallic modulated films for the spectral analysis of X rays¹, the investigation of extremely low diffusivities of one metal in another^{2,3} and the study of phase transformations⁴.

In 1977, Yang *et al.*⁵ prepared a series of multilayer films of the metal pairs Au–Ni

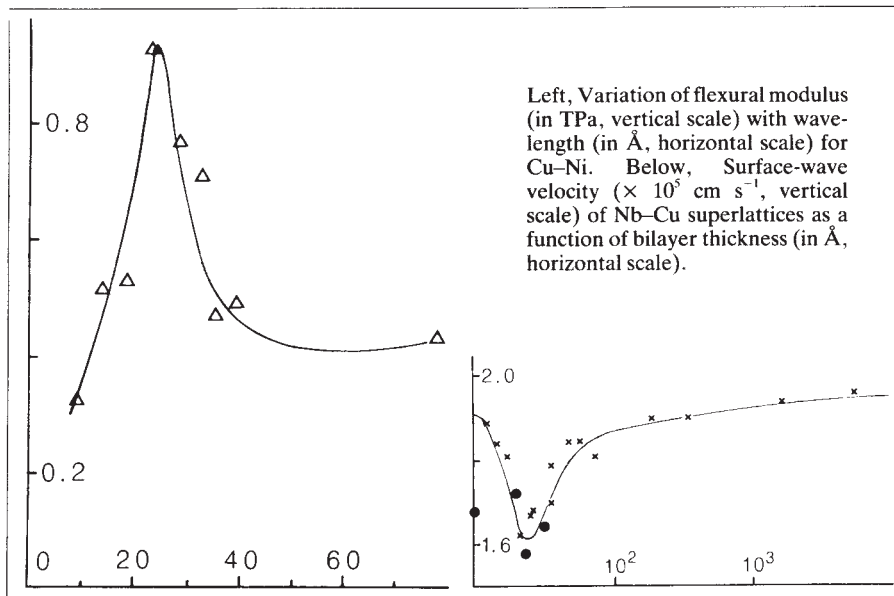
and Cu–Pd. The elastic stiffness of the films was measured by a primitive method that allowed the film to bulge under gas pressure over a small hole; the bulge height is determined as a function of gas pressure. The stiffness thus measured is a biaxial modulus. For Au–Ni the stiffness doubles for a critical thickness near 2 nm; for Cu–Pd, the effect is less pronounced. Recently, Baral *et al.*⁶ used a more conventional, vibrating-reed approach to measure Young's modulus. The figures show their results⁷ for Cu–Ni of constant 50/50 Cu/Ni ratio but varying periodicity and, for comparison, results with measurements of surface-wave velocity for 50/50 Cu–Nb films. The elastic anomaly near 2-nm thickness can evidently be a maximum or a minimum, although taken together, all published

measurements⁸ indicate that the anomaly is usually an enhancement and only rarely (to adopt the latest neologism) a dehancement. In Cu–Au films, however, there is no change. In some of the systems that show a modulus peak at a critical periodicity, there is also a peak of critical stress for plastic deformation.

There are two main rival theories that seek to interpret this 'supermodulus' effect. Clapp⁹ proposed a purely electronic interpretation in which the compositional modulation creates additional Brillouin zones which can touch the Fermi surface and thereby stabilize the modulated structure. If this happens, then an elastic strain will tend to drag the zone and Fermi surface out of contact, which acts in the direction of structural destabilization and thereby enhances the stiffness. (The converse effect is also conceivable.) As various theorists have pointed out (for example, ref. 10), this hypothesis predicts several critical periodicities that should lead to anomalously enhanced (or deenhanced) stiffness.

The alternative view¹¹ is that the supermodulus effect is caused by coherency strains at the metal/metal interfaces, which are postulated to displace atoms so far from their normal sites that they are no longer within the hookean parts of their interaction potential. This approach can also interpret both enhancements and dehancements, but cannot explain the origin of the sharpness of the maximum so easily. For large periodicities, interface dislocations tend to form and release the coherency strains, but this still leaves unexplained the behaviour at periodicities of less than 2 nm. One interesting possibility is that the dehancement of stiffness in Cu–Nb films might result from glass formation by interdiffusion, the so-called solid-state amorphization reaction¹². Cu–Nb would be a reasonable candidate system for this process, though it has not been investigated with the amorphization reaction specifically in view. It is interesting that Lin and Spaepen¹³ find that in Ni–Nb multilayer films, metallic glass is formed spontaneously during film deposition (that is, without anneal) for layer thicknesses smaller than 3 nm.

Very recent work using Cu–(Ni,Fe) multilayers¹⁴ shows the existence of a minimum grain size (~0.1 μm) at 2.6-nm periodicity. At this periodicity, the structure is seen to be defect-free; at smaller periodicities, partial interdiffusion reduces the sharpness of the interfaces and allows grains to coarsen; at periodicities greater than 2.6 nm, interfacial dislocations reduce the elastic coherency strains. These mechanisms can account for the modulus maximum at 2.6 nm. In other recent work¹⁵, elastic moduli are calculated on the basis of the coherency strain model for Cu–Au and Cu–Ni multilayers as a function of periodicity distance, and



roughly agree with experimental values.

The debate between the electronic and coherency strain hypotheses (see refs 8,11) remains wide open, but it does increasingly look as though the coherency strain model will turn out to be correct.

No practical use has been made yet of supermodulus enhancement, but the situation may be about to change. Stiffness is a major concern in the development of improved aircraft alloys, and Al-Li and Al-Li-Be alloys, which are stiff and light, are now receiving concentrated attention. A very new approach is to make massive aluminium alloy pieces, large enough to machine wing spars, by sequential evaporation¹⁶: such alloys have substantially enhanced flow stress as well as stiffness, but there has been no report of a supermodulus effect at critical periodicities. In any case, the periodicities used were 30–40 times larger than in the thin-film work, where peaks in flow stress have been observed at ~2 nm periodicities¹⁵. It remains to be seen whether alloy pieces of useful size can be made with periodicities around 2 nm, and whether an aluminium-based system will show the supermodulus effect. Yahalom¹⁷ has just reported a very promising new approach to making nano-

metre metallic multilayers by pulsed electrodeposition, applied initially to the Cu-Ni system. He specifically suggests applying his method to the production of springs. □

1. Underwood, J.H. & Barbee, T.W. *Nature* **194**, 492 (1981).
2. Greer, A.L. *Scripta Met.* **20**, 457 (1986).
3. Cammarata, R.C. & Greer, A.L. *J. Non-Cryst. Solids* **61** and **62**, 889 (1984).
4. Tsakalakos, T. *Scripta Met.* **20**, 471 (1986).
5. Yang, W.M.C., Tsakalakos, T. & Hilliard, J.E. *J. appl. Phys.* **48**, 876 (1977).
6. Baral, D., Ketterson, J.B. & Hilliard, J.E. *J. appl. Phys.* **57**, 1076 (1985).
7. Kueny, A. *et al. Phys. Rev. Lett.* **48**, 166 (1982).
8. Cammarata, R.C. *Scripta Met.* **20**, 479 (1986).
9. Clapp, P.C. in *Modulated Structure Materials* (ed. Tsakalakos, T.) 465 (Nijhoff, Dordrecht, 1984).
10. Falco, C.M. & Schuller, I.K. in *Synthetic Modulated Structures* (eds Chang, L.L. & Giessen, B.C.) 339 (Academic, London, 1985).
11. Tsakalakos, T. & Jankowski, A.F. *A. Rev. Mat. Sci.* **16**, 293 (1986).
12. Schwarz, R.B. & Johnson, W.L. *Phys. Rev. Lett.* **51**, 415 (1985).
13. Lin, C.J. & Spaepen, F. *Acta Metall.* **34**, 1367 (1986).
14. Jankowski, A.F. & Shewbridge, J.F. *Mat. Lett.* **4**, 313 (1986).
15. Imafuku, M., Sasajima, Y., Yamamoto, R. & Doyama, M. *J. Phys.* **F16**, 823 (1986).
16. Bickerdike, R.L. *et al. Int. J. Rapid Solidification* **1**, 305 (1986).
17. Yahalom, J. in *Materials and Processing Report* Vol. 1 No. 7 (October 1986).

Robert W. Cahn is in the Department of Metallurgy and Materials Science, University of Cambridge, Cambridge CB2 3QZ, UK.

Molecular evolution

Deciphering divergent codes

L. A. Grivell

To evolutionary biologists the protists must appear a motley collection of organisms with diverse characters, defying simple classification, and with many unusual features whose original selective advantage can sometimes only be guessed at. The discovery last year^{1–5} that ciliate protozoa make use of an unusual version of the genetic code, in which UAA and UAG codons specify glutamine instead of stop is entirely in character and it offers molecular biologists a unique opportunity to piece together events in early evolution that may have led to the appearance of the present-day genetic code. Clues to what may have happened are presented in a recent article by Hanyu *et al.*⁶

Until 1981, the genetic code was believed to be the same in all organisms and it was therefore quite reasonable to assume that it had been fixed before the divergence of all life forms, any changes being impossible because they would make nonsense of messages for existing proteins. This belief was shaken by the finding⁷ that different versions of the code are used in the mitochondria of different organisms and that some of these changes must be of relatively recent origin.

The simplest explanation for this extraordinary degree of genetic tolerance within mitochondria seems to be that only a

limited set of proteins are encoded that function entirely within the hydrophobic environment of the inner mitochondrial membrane. Mutations leading to changes in codon assignment, which would be lethal in a more complex genetic system with a wider range of gene products, are apparently acceptable. This may also hold for the simple prokaryote *Mycoplasma capricolum*, in which, as in many mitochondria, the stop codon UGA is used to specify tryptophan (see ref. 5 for a review).

How does the deviant code in the ciliates fit this picture? Are these organisms like mitochondria in being able to absorb changes in their genetic dictionaries, or are we looking at a relic of a primitive genetic code, one which possibly preceded the emergence of the so-called universal code? The first alternative seems unlikely in the light of the complex and definitely eukaryotic characters of the ciliates. The second, in its simplest form, would seem to require that all bacteria, plants, fungi and animals have independently but consistently switched UAA and UAG codons to signal stop. Hanyu *et al.*⁶ suggest a possible solution to this problem.

These authors show the occurrence in *Tetrahymena* of three different transfer RNAs that can attach glutamine. The first has the anticodon UmUG (m indicating

methylation), capable of decoding the two normal glutamine codons CAA and CAG and is probably the counterpart of the glutamine-specific tRNA present in other organisms. The others are unusual and consist of a tRNA^{Gln}_{UmUA} and a tRNA^{Gln}_{CUA}. The first of these recognizes both UAA and UAG in natural messenger RNAs in an *in vitro* protein synthesizing system, while the second recognizes UAG only. Sequence comparison shows that all three tRNAs are related, with tRNA^{Gln}_{UmUA} and tRNA^{Gln}_{UmUG} showing 81 per cent homology and tRNA^{Gln}_{UmUA} and tRNA^{Gln}_{CUA} 95 per cent homology. Hanyu *et al.* therefore conclude that the two unusual glutamine-specific tRNAs evolved from the normal one early in ciliate evolution by gene duplication and divergence, with the UAG-specific tRNA^{Gln}_{CUA} derived from the tRNA^{Gln}_{UmUA} to give a strengthened response to UAG.

They go on to argue that because, apart from mitochondria and *Mycoplasma*, all eubacteria and eukaryotes normally use UAA, UAG and UGA as termination codons, it is more likely that two of these codons failed to specify termination during the evolution of the ciliates, rather than being acquired as stop codons by the other organisms in a process of convergent evolution. They therefore propose that although UAA and UAG were probably stop codons in the archetypal ciliate, they must have been rarely used and that weak suppressor tRNAs for these codons may have existed at a very early stage. Mutations creating UAA and UAG codons within protein-coding genes would not have been lethal for such prociliates. Any subsequent mutation which allowed acquisition of an anticodon capable of perfect recognition of these codons would then result in the fixation of their use as amino-acid codons. A G-to-A transition in the anticodon of a duplicated glutamine-specific tRNA gene is one of the simplest ways in which this could have occurred.

This picture has some attractive aspects. First, it is consistent with the suggestion by Woese⁸ that the primitive genetic code may have been highly ambiguous and error-prone and that even though considerably fewer than the present 20 amino acids may have been encoded, it would have been important to minimize the number of codons not specifying an amino acid. Second, it is consistent with molecular phylogenetic data^{4,9} indicating that the ciliates branched off from the primitive eukaryotic ancestor very early in evolution, quite possibly at a time when protein translation was still tolerant of changes in the genetic code. Finally, it is consistent with more recent patterns of change in evolution, in which gene duplication followed by the acquisition of a new function by one of the duplicated genes is common. Despite these attractive features, however, the process proposed by Hanyu *et al.*⁶ tends to underplay the fact

that many other organisms are able to recognize and translate termination codons which are used either to modulate the synthesis of certain proteins, or to allow the insertion of modified amino acids (for reviews, see refs 10,11). Recognition of termination codons *per se* is thus not unusual. How they are recognized depends on subtleties of tRNA structure, modification or codon context. The usage of UAA and UAG as normal codons in the ciliates may be an extreme form of this recognition process, perhaps one that has developed secondarily as part of the peculiar life style of these organisms.

If, on the other hand, the existence of tRNAs capable of the recognition of stop codons is the relic of an intrinsic property of a primitive and error-prone code, additional variations in the code in other evolutionary groups might be expected, particularly in those subject to similar evolutionary pressures and who branched off the main eukaryotic line either before or around the same time as the ciliates. Based on sequence comparisons of ribosomal RNA (ref. 9), *Dictyostelium*, *Euglena* and the *Trypanosoma* represent even earlier divergence points from the mainstream of eukaryotic descent. But neither *Dictyostelium* nor *T. brucei* shows any deviation from the standard code. Data on *Euglena* and other ancient groups, such as the dinoflagellates, the red algae and *Pelomyxa*, are eagerly awaited.

Finally, a couple of other intriguing questions remain. The first relates to the mechanism by which protein synthesis is terminated. As Hanyu *et al.*⁶ point out, higher eukaryotes seem to use only one release factor for all three termination codons. It will therefore be of interest to see how the *Tetrahymena* factor, which should recognize only UGA, differs from its known eukaryotic counterparts. The other question is whether any organism exists or has existed in which stop codons are not used at all, so that translation is ended by 'run-off' synthesis. It may be worthwhile to delve further among the protists; the answer could be in there somewhere. □

1. Caron, F. & Meyer, E. *Nature* **314**, 185-188 (1985).
2. Preer J.R., Jr, Preer, L.B., Rudman, B.M. & Barnett, A.J. *Nature* **314**, 188-190 (1985).
3. Helftenbein, E. *Nucleic Acids Res.* **13**, 415-432 (1985).
4. Horowitz, S. & Gorovsky, M.A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2452-2455 (1985).
5. Fox, T.D. *Nature* **314**, 132-133 (1985).
6. Hanyu, N., Kuchino, Y., Nishimura, S. & Beier, H. *EMBO J.* **5**, 1307-1311 (1986).
7. Grivell, L.A. (1983) in *Mitochondria 1983: Nucleo-Mitochondrial Interactions* (eds Scheweyn, R.J., Wolf, K. & Kaudewitz, F.) 25-45 (De Gruyter, Berlin, 1983).
8. Woese, C.R. *Proc. natn. Acad. Sci. U.S.A.* **54**, 1546-1552 (1957).
9. Sogin, M.L., Elwood, H.J. & Gunderson, J.H. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1383-1387 (1986).
10. Hatfield, D. *Trends biochem. Sci.* **10**, 201-204 (1985).
11. Chambers, I. *et al.* *EMBO J.* **5**, 1221-1227 (1986).

L.A. Grivell is in the Section for Molecular Biology in the Laboratory of Biochemistry, University of Amsterdam, Kruislaan 318, 1098 SM, Amsterdam, The Netherlands.

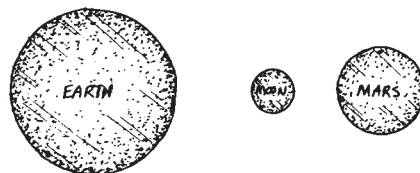
Cosmogony

Moon's birth shocked Earth?

Alan P. Boss

THE most promising explanation of the birth of the Moon is the impact of a Mars-sized body with the Earth during the final phase of planetary accumulation. If such an impact occurred nearly tangentially, the high angular momentum of the Earth-Moon system can be understood, one of the major attractions of this theory. The first detailed computer simulations of such a giant impact have now been completed¹ and can be compared with lunar and geochemical data² to help isolate the manner in which the Moon may have formed.

Thinking on the origin of the Moon has undergone a remarkable transformation in the past few years³. Recent work on the three classical hypotheses of fission, capture and binary accretion has uncovered flaws in their mechanics serious enough to warrant their rejection. But at the same time a relatively new hypothesis has become (not entirely coincidentally) quite



Collision at 10 km s^{-1} of a Mars-sized body with the protoearth would produce a colossal explosion that may have led to lunar formation.

popular. This hypothesis invokes the collision of a Mars-sized or larger protoplanet with the growing protoearth^{4,5}, followed by formation of the Moon in Earth orbit from the debris of the explosion.

When originally proposed a decade ago, the giant impact hypothesis received little attention, largely because giant impacts were thought to be highly improbable. But a new variant of the general theory of terrestrial planet formation predicts that the final phases of formation primarily involved the accumulation of bodies ranging from about the size of the Moon (around one-hundredth the mass of the Earth) to three times the mass of Mars (three-tenths of the Earth)⁶. The appearance of a general theory which predicts the occurrence of Mars-sized impacts on a protoearth of around the present Earth mass with essentially unit probability, coupled with the apparent demise of the competing theories, has encouraged serious efforts at exploring the consequences of the giant impact hypothesis.

Benz and his colleagues¹ use a computer code to model the detailed physics of giant impacts much more accurately than previous efforts⁷. The outcome of a giant impact depends sensitively on impact para-

meter (the offset between the centres of the colliding bodies) and relative velocity. Small impact parameters require high relative velocities to achieve the angular momentum of the Earth-Moon system. Benz *et al.* found that small impact parameters lead to complete disruption of the impactor and to accumulation of this impact ejecta by the protoearth. Their favoured model involves a large impact parameter where the impactor grazes the surface of the protoearth, survives the collision largely undispersed and quickly reforms into a cohesive body. The proto-moon is trapped in Earth orbit by a combination of dissipation of relative velocity during the collision, acceleration by gas pressure gradients in vaporized phases, and gravitational torques with the deformed protoearth. Because the orbit of the proto-moon initially has a perigee within the Roche limit of the Earth (around three Earth radii), the proto-moon may eventually be tidally disrupted, forming a de-volatilized disk of matter which may accumulate into the final Moon. More recent calculations by Benz *et al.* include iron cores in both bodies, and show that the iron core of the impactor can be swallowed by the protoearth, leaving the impactor mantle to form the proto-moon.

These models illustrate the main successes of the hypothesis: giant impacts provide a reasonable mechanism for injecting roughly a lunar mass of matter into Earth orbit, with the right amount of angular momentum, and with the matter depleted in iron and volatiles, as required by chemical analysis of lunar samples. Furthermore, an origin as violent as proposed seems consistent with the large-scale melting required by lunar magma ocean models.

One difficulty with this version of the giant impact hypothesis is now apparent. Ringwood² has re-emphasized the strong similarities in abundance of siderophile elements in the Earth and the Moon. The models of Benz *et al.* imply that the proto-moon formed almost entirely out of the impacting body, thereby requiring the impactor, formed elsewhere in the terrestrial planet zone⁶, to have a mantle very close in composition to that of the Earth. This improbable requirement spoils the beauty of the hypothesis. But there may be a way out.

Melosh and his colleagues^{8,9} have also produced detailed models of giant impacts. Contrary to Benz *et al.*, they find that large-scale vaporization accounts for the formation of a prelunar disk about the Earth. The discrepancy appears to be

caused by different assumptions about the equation of state and by difficult numerical techniques, with the work by Melosh *et al.* probably coming closer to reality. Because vaporization does not discriminate between impactor and impactee, Melosh's prelunar disk is a roughly equal mixture of impactor and protoearth. Making the Moon out of a mixture of impactor and protoearth mantle seems to solve the siderophile problem, because the protoearth composition would also be altered by the addition of the bulk of the impactor, with the final Earth mantle being composed of as little as about 70 per cent protoearth mantle (assuming a three-tenths Earth-mass impactor striking a seven-tenths Earth-mass protoearth), a proportion closer to that of Melosh's prelunar disk.

Clearly many details remain to be investigated and other problems, such as

giant impacts totally melting the Earth and producing a massive crust incompatible with the present Earth, will require explanation. If divining the origin of the Moon were a paying proposition, venture capital firms would certainly be well advised to invest in improved models of giant impacts. □

1. Benz, W. *et al.* *Icarus* **66**, 515 (1986).
2. Ringwood, A. E. *Nature* **322**, 323 (1986).
3. Boss, A. P. *Science* **231**, 341 (1986).
4. Hartmann, W. K. & Davis, D. R. *Icarus* **24**, 504 (1975).
5. Cameron, A. G. W. & Ward, W. R. *Lunar Science Conf. VII*, 120 (1976).
6. Wetherill, G. W. *Science* **228**, 877 (1985).
7. Cameron, A. G. W. *Icarus* **62**, 319 (1985).
8. Melosh, H. J. & Sonett, C. P. in *Origin of the Moon*, 621 (Lunar Planet. Inst 1986).
9. Kipp, M. E. & Melosh, H. J. *Lunar planet. Science Conf. XVII*, 420 (1986).

Alan P. Boss is in the Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road NW, Washington DC 20015, USA.

Fusion energy

Highest temperatures yet

R.S. Pease

DEUTERIUM gas in the Tokamak Fusion Test Reactor at Princeton was heated this summer to a temperature of 230×10^6 K (20 keV), perhaps the highest temperature yet recorded in a laboratory^{1,2}. It lies close to the temperatures planned for fusion reactors, and paves the way for the planned production of megawatts of fusion power in deuterium-tritium experiments in 1989. These promising results bring us closer to the goal of controlled nuclear fusion energy.

What does this experiment amount to, and how does it relate to other attempts? The Tokamak Fusion Test Reactor experimental assembly is a toroidal electric discharge in a strong longitudinal magnetic field. Three other large tokamaks are now operating: the Joint European Torus at Culham, JT-60 at Tokai in Japan and D III-D at San Diego. In the Tokamak Fusion Test Reactor, the torus major radius is 2.4 m and its minor radius is about 1 m. The discharge current in the high-temperature experiments is about 1 MA and lasts about 5 s; this current provides the magnetic field which confines the high-temperature deuterium plasma and provides the necessary thermal insulation; it also provides a relatively small amount of heating which on its own heats the plasma to about 20×10^6 K. The higher temperatures at Princeton were obtained with considerable additional heating provided by injecting into this plasma a stream of energetic neutral deuterium atoms (pre-accelerated as ions to 95 keV and then neutralized).

The additional heating amounted to 17 MW for up to 0.5 s during the latter half of

the current pulse; it raised the ion temperature at the centre of the discharge to 230×10^6 K. This temperature was still rising at the end of the heating period (which is limited by engineering technicalities). The ion number density in the central region (within roughly 20 cm of the centre of the minor cross-section) is of the order of 5×10^{19} m⁻³; and the thermal coupling between the electrons and the ions is relatively slow. The electron temperature is about 70×10^6 K, noticeably lower than that of the ions. The ion temperature was measured by the Doppler broadening of line radiation from nickel and iron impurities (in highly stripped, lithium- and helium-like states), and is consistent with magnetic measurements of the total energy in the plasma (roughly 2 MJ). The energy confinement time, a measure of the thermal insulation obtained by dividing the plasma energy by the input power, is between 0.1 and 0.2 s.

The experimental error of the estimated deuterium ion temperature has not been announced; it is probably of the order of 25 per cent, but until a full exposition of the estimation methods is published, some reservation must remain on the value of the temperature quoted. The temperature planned for in reactors is about the 230×10^6 K (or 20 keV) now claimed. No such uncertainty applies to the deuterium-deuterium fusion power in the plasma, which is unambiguously obtained by measuring the neutron flux emitted from the plasma. This fusion power is reported to be about 10 kW. A substantial but unspecified proportion of this power must be due to direct interaction of the injected 95-

keV deuterons with the plasma deuterons, an enhancing effect explicitly sought in the objectives of the Tokamak Fusion Test Reactor. If a 50/50 mixture of deuterium and tritium had been used in the same conditions, then the fusion power would have exceeded 1 MW. Previously about 0.5 kW of deuterium-deuterium fusion power has been reported from Princeton³.

The achievement of these results primarily arises from the heating power — the largest yet used on a tokamak. As reported both by Princeton and by the Joint European Torus team at a recent symposium⁴, neutral injection is an effective way of heating large tokamaks to the 10 keV range, but is nonetheless accompanied by some degradation of the thermal insulation obtained without it, and a consequent two-fold reduction of the energy confinement time, depending on the power used. An encouraging feature of the Princeton result is that this degradation has not increased at the higher power used. The best heating was obtained using so-called balanced injection: that is, the angular momentum imparted to the plasma by the injected deuterium atoms (which, in the unbalanced case, causes the plasma to rotate at near-sonic speeds) is kept small.

From the point of view of scientific understanding, the key issue of the magnetic thermal insulation remains. The energy losses from the plasma exceed the unavoidable minimum values arising from ion-ion collisions and from the inevitable bremsstrahlung radiation of hot electrons. But in some conditions, and in some other experiments, these minimum values have been approached rather closely. One encouraging indication given by the Princeton result is that a qualitative characteristic of the ideal thermal insulation, the bootstrap current⁵, in which the collisional outward diffusion of the plasma across the magnetic field can by itself sustain the discharge current, has been indicated in some conditions of their reactor.

Overall, the preliminary report from Princeton indicates the achievement of conditions close to those hoped for when construction of the \$300-million apparatus was approved in 1976, and suggests that fusion power in the megawatt range is achievable in the planned deuterium-tritium experiments. The report gives an encouraging foretaste of the results from experiments on all the large tokamak assemblies. □

1. Zarnstorff, M. C. *et al.* *Bull. Am. Phys. Soc.* (in the press).
2. Owens, D. K. *et al.* *Bull. Am. Phys. Soc.* (in the press).
3. Murakami, M. *et al.* *Plasma Phys. cont. Fusion* **28**, 17 (1986).
4. Pease, R. S., Rebut, P.-H. & Bickerton, R. J. *Phil. Trans. R. Soc.* (in the press).
5. Bickerton, R. J. *et al.* *Nature phys. Sci.* **229**, 110 (1971).

R.S. Pease is UKAEA Programme Director for Fusion at the Culham Laboratory, Abingdon, Oxfordshire OX14 3DB, UK.

Mass extinctions

Catastrophic vegetation changes

Margaret E. Collinson

DINOSAURS and ammonites are probably the best-known examples of the many species that became extinct at the Cretaceous-Tertiary (K-T) boundary^{1,2}. Because there are unusually high concentrations of iridium in the clays at the boundary, it is possible that the biotic changes were caused by asteroid impact^{3,4}. Terrestrial phenomena such as vulcanism may, however, have been responsible for the extinctions⁵⁻⁷. Although ecological disruptions of marine biota at this time are widely documented¹⁻⁶, little is known about the influence of the K-T event on terrestrial species⁸. New evidence from fossil land plants (refs 9, 10 and page 148 of this issue¹¹) reveals a widespread ecological trauma, coupled with floral turnover, which occurred at the same time as the marine disruptions.

Late Cretaceous floras were dominated by flowering plants in at least four distinctive floral provinces^{12,13}. In the Northern Hemisphere the Aquilapollenites province includes eastern Siberia and western North America, and the Normapolles province included Europe and eastern North America. In the pollen and spore record several features characterize the K-T event. There was a major floral turnover in the Aquilapollenites province, with extinction of many forms; more restricted turnover in the Normapolles province; but little or no turnover in Southern Hemisphere provinces. A marked increase in the development of modern pollen types occurred in the early Tertiary^{8,12,13}.

The first detailed evidence of the vegetational expression of these floral changes, the 'fern spike', was found during a study of the K-T boundary in the Raton Basin of the western interior of North America¹⁴. Destruction of pre-boundary angiosperm-dominated vegetation was followed by establishment of a fern-dominated, coal-forming mire vegetation. Many Cretaceous flowering plants became extinct, although some persisted across the boundary. Re-establishment of angiosperm dominance followed "within a few centimetres of the boundary"¹⁴ but with very different floral composition.

These events are known to have occurred throughout the western interior of North America and in Canada⁹. Floral change in angiosperm groups falls into three categories: major extinctions but with some species persisting across the boundary; groups showing evolutionary changes between pre- and post-boundary forms; and groups largely unaffected at the boundary. In the first group, different species become extinct in the northern

and southern parts of the region, and macrofloral analysis¹¹ reveals differential extinction in the southern (up to 75 per cent of Cretaceous species) compared with the northern (up to 25 per cent) part of the western interior. The presence of a post-boundary, fern-dominated, coal-forming, mire vegetation was confirmed by the macrofossil analyses.

A recent palynological study¹⁰ from a marine sequence in Japan reveals a fern spike at the same time as the major faunal turnover in marine biota. This study not only extends the geographical range over which the ecological disruption in the Aquilapollenites province has been recognized, but demonstrates, for the first time, that the marine and terrestrial events happened at the same time.

A relationship between leaf form and prevailing conditions has been observed, where vegetation is in equilibrium with climate. Using this relationship, Wolfe and Upchurch¹¹ report in this issue additional aspects of the K-T boundary events in western North America. These include an increase in precipitation across the boundary with the initiation of rain forest environments in the early Tertiary; little overall change between late Cretaceous and early Tertiary temperatures; and a high survival rate of deciduous versus evergreen taxa across the boundary both in northern conifer-dominated and in southern angiosperm-dominated vegetation. An increase in coniferous elements in the post-boundary vegetation occurred in the Montana/Wyoming area of the western interior of North America¹¹ and also in Japan¹⁰. (It may also occur in the western Mediterranean if magnetostratigraphic correlation of the sections studied proves correct¹⁵.)

Vegetational changes across the boundary have been compared with the destruction and renewal of modern vegetation after catastrophic volcanic eruption^{14,16,17}. Replacement climax vegetation may, in such cases, be controlled by the disruptive effects of continued disturbance, without reaching equilibrium with the prevailing climate¹⁸. Thus, palaeoclimatic inferences based on K-T boundary vegetation sequences assume that no subsequent disturbance affected post-boundary vegetation succession. Sedimentological or other evidence is needed to confirm this assumption.

To explain the differential extinction of evergreen taxa, Wolfe and Upchurch¹¹ suggest a 1-2 month low (near 0 °C) temperature change at or near the boundary. The magnitude and duration of such an

excursion is clearly constrained by the high survival rate of large aquatic ectotherms¹¹. But it seems possible that seeds of evergreen taxa survived in the seed bank and re-established when conditions became less hostile.

Despite possible problems with such details, it seems clear that the vegetation did change across the K-T boundary in the Aquilapollenites province⁹⁻¹¹. Angiosperm-dominated forests were destroyed and a fern-dominated, coal-forming mire vegetation became established immediately post-boundary at the same time as marine biotic changes. Turnover in angiosperm floras varied according to taxon and geographical area, but included both major extinctions and persistence more or less unchanged across the boundary. Post-boundary vegetation was therefore different in many ways from that pre-boundary. In some areas conifers increased in abundance and selection for deciduousness had a profound effect on Northern Hemisphere but not Southern Hemisphere vegetation¹¹.

If volcanism in the Indian Deccan Traps did cause the extinctions, as has recently been suggested¹⁹, it is indeed surprising that the Southern Hemisphere provinces show no evidence of the ecological trauma of the Aquilapollenites province. Detailed examination of K-T boundary sections from these provinces is clearly needed. The last appearance of the dinosaurs may well have preceded these vegetational changes⁸, but of the many possible reasons for the extinction of the dinosaurs, those involving a short cold period or enforced changes in herbivore diet bear reinvestigation. □

1. Raup, D.M. & Sepkoski, J.J. *Science* **231**, 833-836 (1986).
2. Raup, D.M. *Science* **231**, 1528-1533 (1986).
3. Alvarez, L.W., Alvarez, W., Asaro, F. & Michel, H.V. *Science* **208**, 1095-1098 (1980).
4. Hsu, K.J. *et al. Science* **216**, 249-256 (1982).
5. Officer, C.B. & Drake, C.L. *Science* **219**, 1383-1390 (1983).
6. Officer, C.B. & Drake, C.L. *Science* **227**, 1161-1167 (1985).
7. Wolbach, W.S., Lewis, R.S. & Anders, E. *Science* **230**, 167-170 (1985).
8. Hickey, L.J. *Nature* **292**, 529-531 (1981).
9. Tschudy, R.H. & Tschudy, B.D. *Geology* **14**, 667-670 (1986).
10. Saito, T., Yamanoi, T. & Kaicho, K. *Nature* **323**, 253-256 (1986).
11. Wolfe, J.A. & Upchurch, G.R. *Nature* **324**, 148-152 (1986).
12. Batten, D.J. in *Fossils & Climate* (ed. Brenchley, P.) 127-164 (Wiley, Chichester, 1984).
13. Muller, J. *Ann. Mo. bot. Gdn.* **71**, 419-443 (1985).
14. Tschudy, R.H., Pilmore, C.L., Orth, C.J., Gilmore, J.S. & Knight, J.D. *Science* **225**, 1030-1032 (1984).
15. Ashrof, A.R. & Erben, H.K. *Palaeontographica* **200**, 111-163 (1986).
16. Leahy, G.D., Spoon, M.D. & Retallack, G.J. *Nature* **318**, 318 (1985).
17. Spicer, R.A., Burnham, R.J., Grant, P.G. & Glickson, H. *Fern. J.* **75**, 1-5 (1985).
18. Cross, A.T. & Taggart, R.E. *Ann. Mo. bot. Gdn.* **69**, 676-734 (1980).
19. Courtillot, V., Besse, J., Vandamme, D., Jaeger, J.-J. & Montigny, R. *C. r. hebdom. Seanc. Acad. Sci. Paris. Ser II* **303**, 863-868 (1986).

Margaret E. Collinson is in the Department of Biology, King's College London (KQC), Kensington Campus, Campden Hill Road, London W8 7AH, UK.

Dating and correlating the French Mousterian

SIR—We note the thermoluminescence dates from Le Moustier¹ with interest and consider them to be valid and informative in the context of that site. However, it is regrettable that in his News and Views piece Paul Mellars² has used the publication of these dates as an opportunity to reinterpret the chronostratigraphy of Combe Grenal and, by inference, the early Upper Pleistocene sequence of south-west France.

Mellars claims that various lines of chronostratigraphic evidence from "key sites" such as Le Moustier and Combe Grenal now converge to suggest that the Mousterian archaeological sequences are successive rather than parallel. Unfortunately, this claim is based on the contentious assumption that data from certain so-called "key" sequences can be simply used to interpret other sites, it omits relevant evidence and it balances perilously on the unjustified correlation of the oxygen isotope record from deep sea core V19-30 with arboreal pollen from Combe Grenal. The correlations suggested by Mellars fail to take into account a conventional radiocarbon date and thermoluminescence dates for Combe Grenal³ and ignore significant stratigraphic non-sequences between layers 55-56 and 35-36 at Combe Grenal and H1 and G4 at Le Moustier. The proposed correlation between the sites would imply an average error of 50 per cent in the available thermoluminescence dates. Although it is clear that the chronostratigraphic interpretation of the Combe Grenal data is not straightforward³, an error of this magnitude seems most unlikely (S.G.E. Bowman, personal communication). Indeed, the dates tend to support the alternative hypothesis of considerable overlap between the Le Moustier and Combe Grenal sequences as suggested by Bordes⁴ and Laville^{5,6}.

As an alternative to the thermoluminescence dates, Mellars compares arboreal pollen data from Combe Grenal with the oxygen isotope sequence. This correlation is problematic in the absence of independent dating to indicate the chronological position and duration of the pollen sequence. Without such dating the comparison of curves that represent ice volume and percentage of arboreal pollen rather than actual temperature is speculative, especially as the latter may vary due to local conditions and taphonomic factors as well as climatic change^{7,8}.

Mellars uses the correlations in his two figures to support his particular view⁹ of the succession of archaeological industries from Ferrassie to Quina and then Mousterian of Acheulian Tradition. This view would seem to be supported by his posi-

tioning of the La Ferrassie hominid from Peyrony's layer D^{10,11} at 60,000-75,000 BP. This is surprising because it requires a hiatus in the order of 30,000-40,000 years between layers D and E, assuming the generally accepted dating of 30,000-35,000 BP for the Lower Perigordian archaeological assemblage in layer E. Further, it does not allow for the possibility that the burials were inserted into layer D and could postdate both the deposits and, in Mellars view, the Ferrassie assemblage. Elsewhere¹¹ the burials have been said to be late Mousterian, dating from Würm II at around 37,000 BP. On the evidence available it would seem unwise to try to arrange the fossil hominid record in line with the model given by Mellars.

Overall, the viewpoint presented seems premature and needs to be considered further in the context of a detailed review of the chronostratigraphic and archaeological evidence. Otherwise, to paraphrase Bordes¹² on the subject, we remain at risk of setting Charlemagne on a motorcycle.

NICK ASHTON

JILL COOK

*Department of Prehistoric
and Romano-British Antiquities,
The British Museum,
London WC1B 3DG, UK*

1. Valladas, H., Geneste, J.M., Joron, J.L. & Chadelle, J.P. *Nature* **322**, 452 (1986).
2. Mellars, P. *Nature* **322**, 410 (1986).
3. Bowman, S.G.E. & Sieveking, G.de G. *Pact* **9**, 253 (1983).
4. Bordes, F. *A Tale of Two Caves* (Harper & Row, New York, 1972).
5. Laville, H. *World Archaeol.* **4**, 323 (1973).
6. Laville, H., Rigaud, J.-P. & Sackett, J.R. *Rockshelters of the Perigord: Geological Stratigraphy and Archaeological Succession* (Academic, New York, 1980).
7. Woillard, G.M. *Nature* **281**, 558 (1979).
8. Turner, C. in *Palynologie Archéologique* (eds Renault-Miskovsky, J., Bui-Thi, M. & Girard, M.) 347 (CNRS, Paris, 1984).
9. Mellars, P. *Proc. Prehist. Soc.* **35**, 134 (1969).
10. Capitan, L. & Peyrony, D. *Rev. Archéol.* **22**, 29 (1912).
11. Delporte, H. (ed.) *Le Grand Abri de la Ferrassie* (Institut de Paléontologie Humaine, Paris, 1984).
12. Bordes, F. *Bull. Soc. Préhist. Française* **78**, 77 (1981).

MELLARS REPLIES—First, I was well aware of the small series of thermoluminescence dates (6 against the 34 for Le Moustier) produced by the British Museum laboratory for the Combe Grenal sequence¹, but these have little if any direct bearing on the central issues. The laboratory itself has already expressed serious reservations over the interpretation of several of these dates¹ and drawn attention to the obvious conflict between the two dates of 105,000 and 113,000 years obtained for the basal levels of the sequence and the totality of the geological, palaeobotanical, faunal and archaeological evidence which points to a *minimum* age for these deposits in the region of 130,000-140,000 BP (= ocean core stage 6, or 'Rissian' in the French terminology)¹⁻³. Dating of a parallel series of samples by the same laboratory from the nearby site of Pech de l'Azé IV has been dismissed by the laboratory as "too young to be acceptable"⁴. For this reason, exact inter-

pretation of the Combe Grenal sequence on the basis of the thermoluminescence dates alone would clearly be hazardous.

What is significant is that 5 of the 6 thermoluminescence dates for the Combe Grenal sequence are substantially earlier than any of the 34 dates for the Le Moustier succession — which is fully consistent with the general conclusion that the greater part of the geological and archaeological sequence at Combe Grenal is entirely earlier than that at Le Moustier. Note that dating of a sample from the upper (but not final) Mousterian levels at Pech de l'Azé IV yielded a date of 19,600 ± 1,600 BP, which is approximately half of the known age of these levels, and demonstrably in error by a factor of at least ten standard deviations on the measured sample age⁴.

Second, I am not quite sure I understand the objections to the proposed correlation between the Combe Grenal sequence and the oxygen isotope record in ocean cores, which is based not on my own interpretations but on the correlations already adopted by Laville and others for the major part of the Combe Grenal succession^{5,6}. The interpretation of the ocean core sequence as a record of climatic fluctuations depends not only on the record of oxygen isotope ratios but also on studies of faunal assemblages in ocean core sediments (as a reflection of sea-surface temperatures)⁷ and on the close correlations now established with palaeobotanical records from terrestrial and offshore sites, particularly the long and well-dated pollen sequence from Grande Pile in north-eastern France^{8,9}. Similarly, the interpretation of the climatic sequence at Combe Grenal depends not only on changing arboreal pollen ratios (though these are the easiest to illustrate in simple graphical form) but on the parallel and mutually consistent records of geology, sedimentology and faunal assemblages^{2,3,10}. The remarkable similarity in all these independent records of climatic change can hardly be dismissed by a passing reference to "local and taphonomic factors" and provides unambiguous support for the climatological/chronological correlations proposed by myself, Laville⁵, Dennell⁶ and others for the Combe Grenal succession.

Third, limited depositional hiatuses should perhaps have been indicated in figures 1 and 2 of my article between layers G and H at Le Moustier and layers 35 and 36 at Combe Grenal — but these episodes have no bearing on the specific correlations proposed between the two sites.

Fourth, I fully accept the implications of a major hiatus between the Mousterian and Upper Palaeolithic levels at La Ferrassie. Similar depositional/erosional hiatuses are well documented at several other cave and rock-shelter sites in south-west France, for example between the

Mousterian and Solutrian levels at Le Placard, and between the Mousterian and Aurignacian levels at Abri Caminade^{2,11}. Positive indications of this hiatus can be seen in recently published sections of the site¹⁴. The suggestion that the Neanderthal burials at La Ferrassie represent later intrusions into the Mousterian levels is intriguing, but difficult to evaluate, 80 years after the original excavations.

Finally, Cook and Ashton make no allowance for the wealth of purely archaeological evidence bearing on the wider problems of Mousterian chronology and inter-correlations between sites. Any attempt to support the stratigraphic correlations of Laville^{2,12} would need to explain why at least 12 sites in south-west France show a consistent sequence of Mousterian of Acheulian tradition above Quina/Ferrassie industries, why Mousterian of Acheulian tradition industries invariably occur at the top of Mousterian sequences in cave and rock shelter sites, and why Mousterian of Acheulian tradition industries are entirely lacking from all except

the uppermost 4 of the 55 levels of Mousterian occupation at Combe Grenal¹³. No explanation has so far been offered as to how these observations can be reconciled with the hypothesis of a close synchronism between the archaeological sequences at Le Moustier and Combe Grenal.

PAUL MELLARS

Department of Archaeology,
Downing Street,
Cambridge CB2 3DZ, UK

1. Bowman, S.G. & Sieveking, G. de G. *Pact* 9, 253 (1983).
2. Laville, H., Rigaud, J.-P. & Sackett, J.R. *Rock Shelters of the Perigord: Geological Stratigraphy and Archaeological Succession* (Academic, New York, 1980).
3. Bordes, F., Laville, H. & Paquereau, M.-M. *Actes de la Société Linnéenne de Bordeaux* 103B, 1-19 (1966).
4. Bowman, S.G.E. et al. *Pact* 6, 362 (1982).
5. Laville, H. et al. *Bull. Inst. Géol. Bassin Aquitaine* 34, 219 (1983).
6. Dennell, R. *Nature* 301, 199 (1983).
7. Saccetta, C. et al. *Quat. Res.* 3, 110 (1973).
8. Woillard, G. & Mook, W.G. *Science* 215, 159 (1982).
9. Turon, J.-L. *Nature* 309, 673 (1984).
10. Bordes, F. & Prat, F. *L'Anthropologie* 69, 31 (1965).
11. de Lumley, H. (ed.) *La Préhistoire Française* (CNRS, Paris, 1976).
12. Laville, H. *World Archaeol.* 4, 323 (1973).
13. Mellars, P.A. *Proc. Prehist. Soc.* 35, 134 (1969).
14. Delporte, H. (ed.) *Le Grand Abri de la Ferrassie* (Institut de Paléontologie Humaine, Paris, 1984).

What can AIDS virus codon usage tell us?

SIR—Macromolecular sequence comparison is extraordinarily useful in elucidating evolutionary relationships, largely because divergence rates are approximately constant, and the chance of convergence is very small. The close evolutionary relationship of the human immunodeficiency virus (HIV) and visna virus has been well characterized by this approach^{1,2}. Grantham, largely responsible for demonstrating that patterns of codon choice are species are species specific³, and Perrin now suggest that comparison of such pat-

terns in HIV and visna may provide a useful approximate evolutionary distance⁴.

Decomposition of a DNA sequence into a codon usage table loses a great deal of information — comparisons of the intact sequences of homologous genes and proteins must be more informative than comparisons of codon choice. In fact, the distance measure used by Grantham and Perrin (how many of 18 amino acids have the same most-preferred codon in two codon usage tables) can yield some strange results. For example, among the viruses in Table 1, it is thought that Moloney murine leukaemia virus (MMLV, another retrovirus) is the most closely re-

lated to HIV, and yet HIV codon usage is most similar to that in the influenza virus, and even cauliflower mosaic virus (CAMV) is more similar to HIV than is MMLV. Also, a compilation of 66 mouse genes⁵ yields the same preferred codons as the human (total) for each of the 18 amino acids, and yet differs from 60 rat genes⁶ for two amino acids. Thus, unlike divergence in protein and DNA sequences, differentiation at the level of codon usage is clearly not linear with time. It is also highly susceptible to convergence.

Furthermore, it is clear that there is considerable heterogeneity in codon choice among genes within species. In *Escherichia coli* and yeast (the two best studied organisms) groups of genes with high and low codon bias can be identified^{3,6,7} that differ somewhat in their preferred codons (Table 1). In mammals there are also large differences among genes in the direction of codon preference⁸. These differences seem to be related to local genomic G+C content⁸, and can yield very different sets of preferred codons.

Thus, when Grantham and Perrin use their comparison to estimate the extent to which viruses imitate the codon preference of their hosts⁴, and point out that codon usage in HIV genes is different to that in a compilation of all human (excluding immune system) genes, this is somewhat misleading. In fact, HIV genes show considerable similarity in codon choice to a group of human genes with low G+C content. Comparing species totals; HIV codon usage shows more similarity to that in *Bacillus subtilis* and yeast, than to that in the compilation of human genes.

On the positive side, it is possible that the differences among mammalian genes in pattern of codon usage may partly reflect the diverse tissues in which those genes are expressed. One investigation of this possibility⁹ gave negative results, but examined very few genes. It may be that codon usage in HIV genes reflects adaptation to expression in particular tissues, in which case the report⁴ that there is some similarity in this respect between HIV genes and T-cell receptor genes would be rather interesting.

PAUL M. SHARP

Department of Genetics,
Trinity College,
Dublin 2, Ireland

Table 1 Similarities of codon choice between organisms

	(N)	HIV	E LO	E HI	Baci	Y LO	Y HI	H AT	H CG
Human (total)	(135)	2	8	8	1	2	7	3	14
Human G+C rich	(8)	0	8	10	1	0	7	1	
Human A+T rich	(8)	12	8	5	13	14	7		
Yeast HI	(38)	5	2	11	6	11			
Yeast LO	(66)	11	6	6	13				
<i>Bacillus</i>	(21)	11	9	7					
<i>E.coli</i> HI	(42)	2	8				5	(3)	MMLV
<i>E.coli</i> LO	(58)	6					1	(28)	AD2
				11	1	5	13	(8)	Flu
			10	7	5	8	8	(2)	CAMV
			Flu	SV40	AD2	MMLV	HIV		

Values are the number of amino acids (out of 18) where the two data sets show the same preferred codon. N is the number of genes in each data set. Yeast, *Saccharomyces cerevisiae*; *Bacillus*, *B. subtilis*. HI and LO (for *E.coli* and yeast) indicate groups of genes identified as having high and low codon bias. Human GC-rich: α -actin, apolipoprotein A1, β -chorionic gonadotropin, α -globin, ζ -globin, metallothionein II, insulin, α -tubulin. Human AT-rich: albumin, α -amylase (pancreatic and salivary), factor IX, γ -fibrinogen, γ -interferon, HPRT, parathyroid hormone. Viruses: HIV, 5 genes (*gag*, *pol*, *sor*, *env*, *P27*) of ARV-2; Moloney murine leukaemia virus; Flu, influenza virus isolate B/Lee/40; SV40, simian polyoma; CAMV, cauliflower mosaic virus; AD2, adenovirus 2. All codon usage data was taken from ref.5, except: yeast⁶, *Bacillus*¹⁰, *E.coli*⁷ and HIV (calculated from GenBank).

1. Sonigo, P. et al. *Cell* 42, 369-382 (1985).
2. Yasunaga, T., Sagata, N. & Ikawa, Y. *FEBS Lett.* 199, 145-150 (1986).
3. Grantham, R. et al. *Nucleic Acids Res.* 8, r49-r62 (1980); 9, r43-r47 (1981).
4. Grantham, R. & Perrin, P. *Nature* 319, 727-728 (1986).
5. Maruyama, T., Gojobori, T., Aota, S. & Ikemura, T. *Nucleic Acids Res.* 14, r151-r197 (1986).
6. Sharp, P.M., Tuohy, T.M.F. & Mosurski, K.R. *Nucleic Acids Res.* 14, 5125-5143 (1986).
7. Sharp, P.M. & Li, W.-H. *Nucleic Acids Res.* 14, (in the press).
8. Ikemura, T. *Molec. Biol. Evol.* 2, 13-14 (1985).
9. Hastings, K.E.M. & Emerson, C.P. *J. molec. Evol.* 19, 214-218 (1983).
10. Ogasawara, N. *Gene* 40, 145-150 (1985).

Colonization of space: destiny or folly?

Von R. Eshleman

The US National Commission on Space has grandiose plans for the colonization of the Solar System. Has the view that unmanned missions may bring much greater benefit to mankind been passed over?

THE presidentially-appointed National Commission on Space has recently released an expansive agenda for the US space programme in the first quarter of the twenty-first century (*Pioneering the Space Frontier*, Bantam Books, Inc., New York, 1986). With space enterprise, exploration and science as components, the overall rationale behind this vision of the future is founded on an assumption: that the human species is destined, within the foreseeable future, to colonize space.

But this assumption is not critically examined. It is offered as an act of faith, as though it were rendered self-evident by references to Columbus and 'The New World', the settlement of the American West, and the pioneering spirit that has served us so well in the past.

These are indeed powerful images, but do they necessarily make space colonization a human imperative? I do not question whether it could be done so much as to wonder why it is assumed that it should be done. I would worry about unresolved difficulties, but my main concern is the danger of precluding alternative possibilities of greater promise.

The commission envisages that by the year 2000, the United States would have an Earth Spaceport, with crew, for storage, fueling, repair, supply, maintenance, and launch of transfer and orbital manoeuvring vehicles for access to more distant locales. By the year 2025 there would be: three additional spaceports with human crews, including one near Mars; staffed spaceships continually cycling between the orbits of Mars and Earth; manufacturing at a permanent human base on the Moon; astronaut outposts on Phobos and Deimos, the small moons of Mars; and an exploratory base on the surface of Mars which would become a fully-staffed permanent base within a few years. It is stated that this 'Highway to Space' and 'Bridge Between Worlds' would provide the infrastructure for space applications and science, while serving as a focal point for international cooperation and advances in technology. And it would initiate the colonization of the Solar System.

Will this proposal for pioneering the space frontier prove to be prophetic or visionary? It is intended to help influence future events so as to become the former, but current events and their origins suggest that it could well be the latter. It is my opinion that the commission's proposals, if initiated, would lead to a continuation of

the fundamental failure of the US space programme to use national resources most effectively for progress at the frontier of space. The crux of the problem, I submit, is an inappropriate mix of the relative influences of our space ends and means . . . and sometimes dreams.

In the report, for example, it is the dream of colonization that determines the proposed infrastructure of spaceports, cycling spaceships and lunar and martian bases. But these are simply means; the ends are primarily the exploration and commercial and scientific use of the space environment. Is there not a danger that the means will become ends in themselves? The question is not even asked the other way around. If all manner of space manufacturing, Earth remote sensing, space communications, planetary exploration, and space science laboratories and observatories are indeed our main ends, by what means may they best be realized? This inevitably brings up the decades-old debate concerning the relative importance and roles of manned and unmanned space endeavours. In the commission's view, the debate is finished. Human crews, settlers, and vacationers belong wherever they conceivably can go. In this scenario, the next frontier will be conquered as before, with a wave of human pioneers.

Ultimate frontier

But mankind could expand into the ultimate frontier of space in another way. Never before have we had the ability to observe, explore, utilize, examine, monitor, manipulate, prospect, mine and manufacture, or in short pioneer and operate at a new physical frontier, without the necessity of a physical human presence at that frontier. In a way, humanity as a whole could experience its continuous expansion into thousands of locales at the frontier of space, with a multitude of its doers (like the hand of Viking 1 as it clutched a sample from an ancient plain of Mars) and sensors (like the eyes of Jason Jr as it descended the grand staircase of the Titanic or the magnetometers of Voyager 2 as they charted the bizarre field of Uranus) on the ends of thousands to billions of kilometre-long radio nerve fibres and commandable sinews. Remote human monitoring and control of space manufacturing and delicate experimentation could also be accomplished.

The means to secure these ends would be low-cost unmanned launch and transfer

vehicles and new levels of capabilities in robotics, communications and control, adaptive and expert systems, and many related technologies. Our collective space dreams could encompass the endless possibilities for international cooperation in obtaining new knowledge, products and resources from space in this way, and the endless potential of applying them to better the human condition on Earth. Such dreams include an automation revolution to rival the industrial revolution in significance, and new levels of understanding of climatic effects like carbon-dioxide warming, nuclear winter and acid rain from remotely controlled examinations of related phenomena on other worlds. Concerns about national prestige and demonstrations of technological prowess, as with Apollo, or dreams about solving international tensions by a joint symbolic venture in space, such as a human mission to Mars, do not appear to me to be useful guides for progress.

The current hiatus in American access to space due to the Challenger disaster should be used to re-appraise our goals and the methods for their realization. It is clear that the present problems would have surfaced in some way, even if the tragic accident had not occurred. Although many observers have cheered the decision by President Reagan to replace Challenger, some seem to have missed the significance of the accompanying announcements that the United States will phase out commercial and foreign payloads for the shuttle fleet, and will encourage the commercialization of launch services after the 15-year delay that was enforced by nationally-subsidized competition. This stands as a tacit admission that we have failed to realize the fundamental rationale for the space shuttle programme, namely, that of providing low-cost access to space for all US and many foreign users.

The shuttle experience stands as one example of the kinds of things that could go wrong if we try to bypass incremental, well-funded, long-term progress based on carefully defined goals and feasible, cost-effective means tailored to those ends. I submit that space colonization is a dream whose time has not yet come, and that the benefits of the space frontier beckon us to realize them in the most direct and effective ways possible. □

Von R. Eshleman is Director of the Center for Radar Astronomy, DB221, Stanford University, Stanford, California 94305, USA.

The quark plasma

H. Satz

Fakultät für Physik, Universität Bielefeld, Postfach 8640, D-4800 Bielefeld 1, FRG
and Physics Department, Brookhaven National Laboratory, Upton, Long Island, New York 11973, USA

Because nucleons are bound states of quarks, nuclear matter should at sufficiently high density become a quark plasma. Assuming quantum chromodynamics as the basic theory of nuclear forces, it is possible to derive strong-interaction thermodynamics and study the nature of this transition. Predictions of thermodynamic variables are obtained by computer simulation; they provide the first theoretical information on the phase structure of strongly interacting matter. The experimental search for the quark plasma is beginning this year at Brookhaven and at CERN. Energetic nuclear collisions could produce sufficiently dense and long-lived strongly interacting systems to allow a first look at the predicted new state of matter.

IN the past two decades our picture of an elementary particle has changed significantly. Today we consider protons, neutrons and all the other 'elementary' particles in the realm of nuclear forces as composite: as bound states of quarks. Elementary particles are no longer the smallest units of matter; they, in turn, consist of quarks. The binding force between quarks is quite different, however, from the forces usually encountered in physics: at short distances it becomes so weak that the quarks appear to be free, but with increasing distance of separation, it grows without limit. Thus it is not possible to split an elementary particle into its quark constituents—this would require an infinitely large force. Elementary particles remain the smallest units of matter that can exist by themselves, individually; quarks can only exist together with other quarks or antiquarks.

This modified view of the elementary particle concept has in turn had a profound effect on our picture of nuclear matter at high density. If nucleons (protons and neutrons) were elementary hard spheres, then a densely packed system of such objects would be the limiting state of matter (see Fig. 1). For composite nucleons, however, interpenetration will occur if the density is increased enough, and eventually each quark will find very many others in its immediate neighbourhood. Because it has no way to remember which of these were its partners in the low-density nucleonic state, the concept of a nucleon and of nuclear matter has become meaningless here. Thus it is expected that at a high enough density a new state of matter will appear: the quark plasma—a system whose basic constituents are unbound quarks.

Before considering in more detail how quark binding can be dissolved in dense matter, let us get some idea of the densities involved. Dense atomic solids contain about 10^{23} nucleons cm^{-3} ; nuclear matter, as it exists in the interior of neutron stars, has about 10^{38} nucleons cm^{-3} , and is thus 10^{15} times denser than matter in the world around us. What is the effect of further compression in this regime of extreme density?

I have already noted that the strength of the binding force between quarks increases indefinitely with their separation, making the 'ionization' of a nucleon into its quark constituents impossible: the quarks are 'confined'. Why then should they become deconfined in sufficiently dense matter? The answer to this question is found in charge screening, which already occurs in atomic solids. The force between the charged partners of a bound state is considerably modified, if this bound state is placed in an environment of many other such objects. The Coulomb potential between two electric charges e , separated by a distance r , has in the vacuum the form e^2/r ; in the presence of many other charges it becomes subject to Debye screening, $e^2/r \rightarrow (e^2/r) \exp(-r/r_D)$, where the screening radius r_D is inversely proportional to the overall charge density of the system: the

more other charges present in the environment, the shorter the range of the force between two given charges. If in atomic matter the Debye radius r_D becomes less than the atomic radius r_A , then the binding force between electron and nucleus is effectively screened, and the electron becomes 'free'. For atomic systems, an increase in density thus results in an insulator-conductor transition^{1,2}: the binding between charges is dissolved in the presence of enough other charges.

Screening, in contrast to ionization, is thus a short-range mechanism to dissolve bound states, and for this reason we expect it to apply also to quark binding. I shall return shortly to the theoretical basis of the force between quarks; note here only that the interaction is again based on a charge, the 'colour' charge, which in strong-interaction physics plays the role of the electromagnetic charge in atomic physics. Nucleons are colour-neutral bound states of colour-charged quarks, and hence nuclear matter corresponds to a colour insulator. With increasing density, colour screening should set in and lead to the transition from colour insulator to colour conductor: to the quark plasma. The predicted new state of matter thus is a macroscopic system which is conductive to the basic charge of the strong interaction.

The theoretical analysis of strongly interacting matter is based on the fundamental theory for the interaction of quarks, quantum chromodynamics. I shall survey the present state of this analysis in the next section. The experimental investigation of these phenomena is the main aim of forthcoming studies of

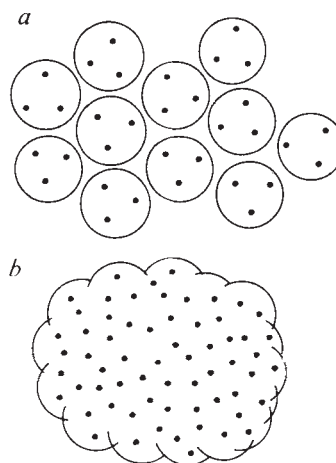


Fig. 1 A system of nucleons (a) as nuclear matter and (b) as quark plasma.

ultrarelativistic collisions of heavy nuclei; this will be taken up in the section after next.

Statistical quantum chromodynamics

Quantum chromodynamics (QCD) describes the interaction of quarks in much the same way as quantum electrodynamics (QED) does for electrons. Spinorial matter fields ψ —here quarks; electrons in QED—interact through massless vector gauge fields A ; here gluons take the place of photons. Just as in QED, the interaction in QCD is determined by an intrinsic charge, the 'colour'; each quark can be in one of three different colour charge states. In contrast to QED, however, the gauge fields here also carry a charge: each gluon can be in one of eight colour states. As a result, the gluons, as well as the quarks, interact with each other, and it is in fact this gauge-field interaction that is the crucial feature of QCD. Gluonic lines of force between two quarks now attract each other, and this turns the three-dimensional Laplace equation, with the Coulomb solution $V(r) \sim 1/r$, into a one-dimensional form, with $V(r) \sim |r|$ as 'confining' solution.

The Lagrangian density of QCD has the form³⁻⁵

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^a F_{\mu\nu}^a - \sum_f \bar{\psi}_f^a (i - g)_{\alpha\beta} \gamma_\mu \psi_f^a \quad (1)$$

with

$$F_{\mu\nu}^a = (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a - gf_{bc}^a A_\mu^b A_\nu^c) \quad (2)$$

Here $a, b, c = 1, \dots, 8$, and $\alpha, \beta = 1, 2, 3$ denote the gluon and quark colour charges, respectively; they are based on SU(3) as the colour gauge group, whose structure constants are f_{bc}^a . The explicit summation in equation (1) runs over the possible different species ('flavours') f of quarks; I restrict myself here to the (practically massless) u and d quarks, which suffice to construct all non-strange nucleons and mesons.

Note again that it is the last term in equation (2), and hence the non-abelian nature of the gauge group⁶, which constitutes the crucial feature of the theory; for $f_{bc}^a = 0$, equations (1) and (2) have the same structure as QED.

The Lagrangian (1)/(2) contains a single, dimensionless coupling g . As a result, QCD cannot predict any dimensional observable, but only the dimensionless ratios of observables. In many ways, such a description comes close to an ideal physical theory: it is closed in itself, and contains no outside parameters. As shown below, QCD thus predicts for instance the transition temperature from nuclear matter to quark plasma in units of the nucleon mass, and this mass is itself calculated within the theory as the lowest bound-state energy of three massless quarks.

Given the Lagrangian density (1)/(2), we have in principle all the dynamical input necessary for strong interaction thermodynamics. From $\mathcal{L}$, we must determine the Hamiltonian H , and then construct in the usual way the partition function

$$Z(T, V) = \text{Tr}(\exp(-H/T)) \quad (3)$$

for a thermodynamic system at temperature T and within a spatial volume V ; the trace in equation (3) runs over all microscopic physical states. From Z , we obtain by differentiation the desired thermodynamic observables, getting for instance $\varepsilon = (T^2/V)(\partial \ln Z / \partial T)_V$ for the energy density, and $P = (\partial \ln Z / \partial V)_T$ for the pressure. But how is it possible to carry out the actual evaluation of Z ? The perturbative approach familiar from QED could be applied here as well for the regime in which the interactions are weak—that is, at very high density. But the central theme of statistical QCD, the transition from the strongly coupled nuclear matter to the weakly coupled plasma, is by nature non-perturbative and cannot be addressed in such a fashion.

The only approach known at present that permits us to study strongly interacting matter at all densities is the lattice formulation introduced by Wilson⁷. Replacing the space-time continuum, on which the fields ψ and A are defined, with a discrete

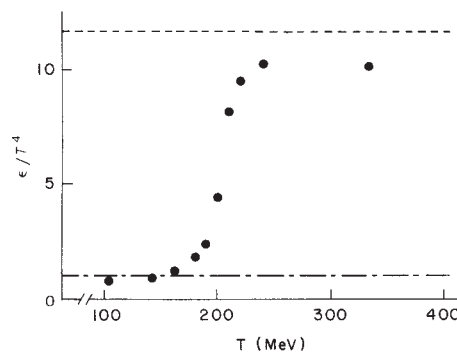


Fig. 2 The energy density of strongly interacting matter as a function of the temperature compared with the energy density of the ideal quark-gluon plasma (---) and of the ideal pion gas (- - - -). (From ref. 12.)

and finite set of lattice points, one obtains a form of the partition function that is divergence-free and, moreover, structurally very similar to that of spin systems in statistical mechanics. It is therefore possible to use the computer simulation techniques developed for such spin systems in the evaluation of lattice QCD⁸. Renormalization group theory, which studies the scale-dependence of statistical systems, then permits us to recover the continuum limit (vanishing lattice spacing) of the results thus obtained. In the past few years this method of evaluation has been applied extensively and with considerable success in the study of the phase structure of strongly interacting matter (for a recent survey, see ref. 9).

After taking a brief look at the computer simulation of statistical systems¹⁰, I shall here go on directly to the results obtained so far.

The analytical mathematics available today has been developed to a large extent for isolated systems; even very simple many-body systems with interaction cannot generally be solved, once the interaction really connects the whole system and is not perturbatively reducible to small parts of it. As an example, recall that the three-dimensional Ising model is not analytically solved. In recent years, with the advent of large-scale computers, it has become possible to simulate such systems: the given system is reproduced on the computer in a specific microscopic configuration. In reproducing all possible configurations, it would then be possible to evaluate $Z(T, V)$ exactly for systems of finite size. Because the three-dimensional Ising model— N^3 spins pointing either up or down—already leads to 2^{N^3} configurations, it is evident that for meaningful systems such a procedure is not feasible, even on the largest computers. Therefore, in statistical physics Monte Carlo methods have been developed to generate 'typical' configurations¹¹, and, averaging over a sufficiently large number of these for the simulated system, it is possible to 'measure' the thermodynamic quantities of interest.

The density of strongly interacting matter can be increased in two distinct ways: by compression, or by heating, which owing to the relativistic energy range involved leads to particle production and hence a higher density. So far most of the results of statistical QCD deal with the latter of these two cases, with 'hot mesonic matter' of vanishing baryon number density. The other extreme, the compression of 'cold nuclear matter', is now the subject of much study; the computer simulation of this system is technically more difficult.

Beginning with matter of vanishing baryon number density, the energy density of a non-interacting gas of massless quarks and gluons is

$$\varepsilon_P = (37\pi^2/30)T^4 \approx 12T^4 \quad (4)$$

Just as in the Stefan-Boltzmann form for the photon gas, the numerical factor in equation (4) is determined by the number

of degrees of freedom of the constituents involved: their spins, colours and flavours. Similarly, a gas of non-interacting π mesons becomes, if we neglect the meson mass, $\varepsilon_\pi = (3\pi^2/30)T^4 \approx T^4$, taking into account the three possible charge states of the π . In the low-temperature region, where quarks are confined, we might expect values of ε in this range, whereas at high enough temperatures the plasma form (4) should become correct. The energy density obtained by computer simulation of statistical QCD¹² is shown in Fig. 2; there is a very rapid transition from the mesonic regime to the plasma. The transition temperature—converted to physical units through comparison with the mass of the nucleon or the ρ meson—is around 200 MeV, which in turn means that an energy density of at least 2.5 GeV fm^{-3} is necessary to create a quark-gluon plasma.

Before we turn to nuclear matter, a further aspect related to the deconfinement of quarks should be noted. An electron in a conductor has an 'effective' mass different from that in a vacuum or in an atom. This mass shift is the consequence of collective effects of all the positive ions and other electrons present. Similarly, we can consider an 'effective' quark mass, which in the confined phase is around 300 MeV (corresponding to the three constituent quarks of a nucleon); in the weakly interacting plasma it vanishes. This shift in the effective quark mass need not exactly coincide with deconfinement. In principle it may occur later, which would lead us to a three-phase structure for strongly interacting matter: hadronic matter, consisting of mesons and nucleons; a constituent quark plasma, consisting of massive but deconfined quarks; and finally the quark-gluon plasma of massless quarks and gluons. For systems of vanishing baryon number density, however, statistical QCD has shown that the two phenomena indeed do occur at the same point^{12,13}.

Systems of non-vanishing baryon number density—'nuclear' or 'baryonic' matter—are now a subject of considerable attention. The partition function depends in this case on a further variable: the baryonic chemical potential μ , which through $n_B = (T/V)(\partial \ln Z(T, V, \mu)/\partial \mu)$ determines the overall baryon number density of the system (the number of baryons minus that of antibaryons, per unit volume). So far, numerical studies have been made only for rather small values of μ ^{14,15}. The resulting phase diagram is shown in Fig. 3; up to this point, deconfinement and quark mass shift still coincide. Whether that will remain true even at very low temperatures, or whether the mentioned additional phase, the constituent quark plasma, makes its appearance there, remains to be seen.

Nuclear collisions at high energies

The thermodynamics of strongly interacting matter is of importance to cosmology. Very soon after the Big Bang, the Universe must have been a hot, dense quark-gluon plasma; its rapid expansion presumably reduced the density such that after about 10^{-3} s the transition to the hadronic Universe occurred. Only from that point on was there a physical vacuum.

In recent years the possibility of recreating this cosmological process in the laboratory has come within reach¹⁶. The collision of two heavy nuclei at high enough energies should produce a 'little bang', providing a bubble of strongly interacting matter sufficiently large and of sufficiently long lifetime to allow the application of statistical QCD; that is, it should form the quark plasma, and make possible the study of the transition to hadronic matter. What is the physical basis for this?

When two nucleons collide at high energy, as for example in one of the large proton accelerators at CERN (Geneva) or Fermilab (Illinois), they almost always just pass through each other. In the process the nucleons are excited, but in addition they leave a 'condensation trail' of deposited energy in the region of space they passed through (see Fig. 4). This energy is rapidly transformed into hadrons, the secondary particles usually produced in such collisions. From the analysis of proton-proton collision data, it is known that the deposition of energy at high enough energies is about $1/3 \text{ GeV fm}^{-3}$. It does not increase

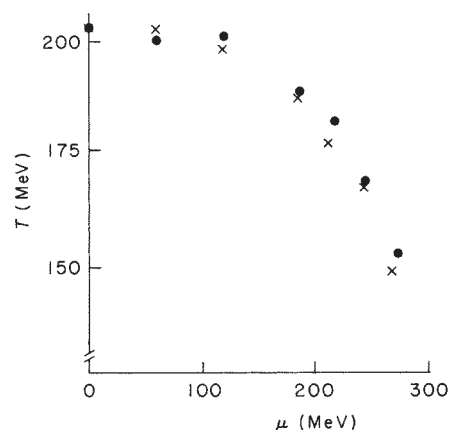


Fig. 3 The dependence of the deconfinement temperature (x) and the chiral-symmetry restoration temperature (●) on the baryonic chemical potential μ ; increasing μ means increasing the baryon number density. (From ref. 15.)

much more with a further increase in collision energy, which only makes the condensation trail longer.

If instead of the nucleons, two heavy nuclei collide (for simplicity, both with mass number A), a multiple superposition of the phenomenon just described is obtained, so that a given region of the condensation trail now receives a much larger deposition of energy¹⁷. Simple geometrical arguments lead to an energy density of at least $\varepsilon \sim A^{1/3} \text{ GeV fm}^{-3}$; this means that an energetic collision of two ^{238}U nuclei provides an average deposition of about 6 GeV fm^{-3} . As seen in the previous section, this is well above what is needed for plasma formation.

Processes of this type can also be simulated on large computers, using simple phenomenological interaction models as input. One such stimulation (done by D. Strottman (Los Alamos), based on a model he developed with A. S. Goldhaber (Stony Brook)) is shown in Fig. 5; we see clearly the formation of the condensation trail after the nucleons have penetrated each other. The process shown here is the collision of a ^{16}O nucleus with a ^{238}U target, at an incident energy of 100 GeV per nucleon. Such reactions will be the first to be studied in forthcoming experiments at CERN (Geneva); I shall return to the experimental situation shortly.

Nuclear collisions should thus produce bubbles of rather high energy density, and of a volume $V_A \sim A \text{ fm}^3$; although not really macroscopic, V_A is, for large A at least, much bigger than the typical hadron volume of about 1 or 2 fm^3 . It is therefore hoped that such a bubble represents a system to which the equation of state obtained in statistical QCD may be applied; if the energy

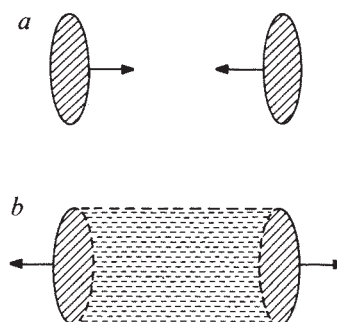


Fig. 4 High-energy collision of two nucleons: (a) just before, (b) just after the collision.

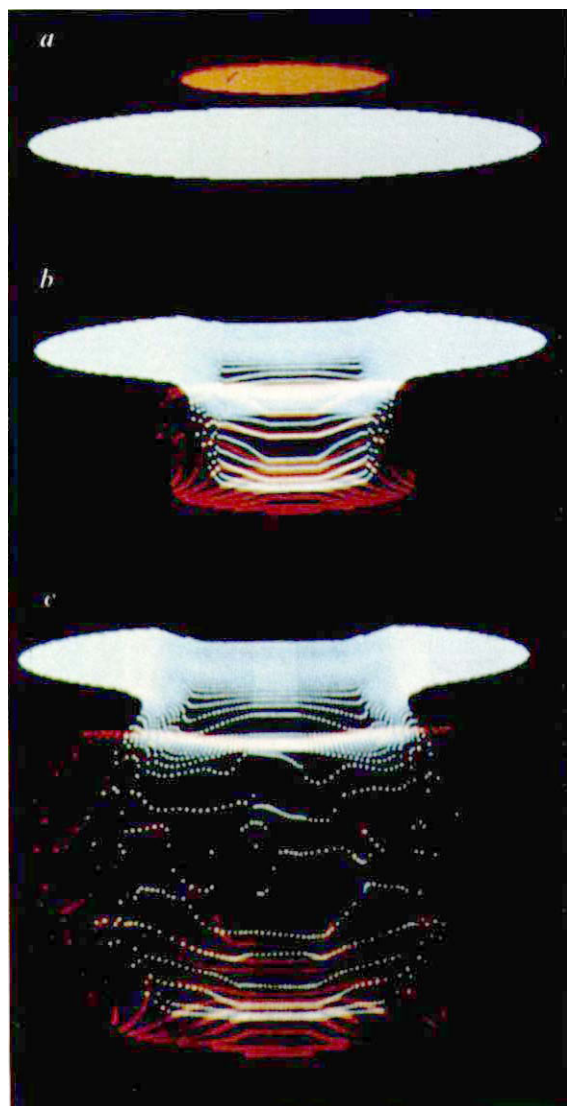


Fig. 5 Successive stages of a high-energy nuclear collision: ^{16}O at 100 GeV per nucleon on ^{238}U ; (a) at approximately 10^{-24} s before the collision, (b) at 2×10^{-23} s and (c) at 10^{-22} s after the collision. (From the computer simulation by D. Strottman.)

deposit is high enough, the initial temperature will be high, and the bubble will contain quark matter.

Each plasma bubble will now expand quite rapidly and in doing so cool down to the deconfinement transition temperature. Hadrons—predominantly mesons—will form, and after still further expansion, at the energy density of about $1/3 \text{ GeV fm}^{-3}$, they will cease to interact and form the secondary particles produced in the collision. The plasma expansion has been a subject of much recent work¹⁸; it is usual to apply the hydrodynamic flow equations, using an equation of state of strongly interacting matter (ideally that obtained from statistical QCD) as input.

Given this model for plasma formation and transition in nuclear collisions, how can it be tested experimentally? What are the signatures for the creation of a new state of matter in such interactions? What has emerged so far is a number of different indications, which, if all observed together, would make plasma formation the most likely explanation.

At the time when any measurement is made, the 'little bang'—provided it occurred—is over. Hence one thing to look for is a signal from the initial plasma phase that is not affected by the

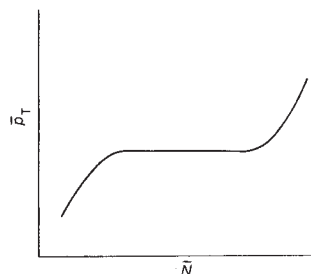


Fig. 6 Schematic dependence of the average transverse momentum $\bar{p}_T$ of hadronic secondary particles on the average number $\bar{N}$ of such secondary particles, as expected for energetic nuclear collisions leading to a first-order deconfinement–confinement transition. (From ref. 19.)

subsequent transition. If quarks in the plasma, by collision or annihilation, emit photons, then these, because of the much weaker electromagnetic interaction, will generally escape from the system without further interactions. The spectra of photons, real or virtual (as e^+e^- pairs), should thus tell us something about the early stage of the produced system.

Because the number of produced secondary hadrons is directly related to the initial energy density, the multiplicity in a given collision may serve as an indicator for ϵ . In the condensation trail we find bubbles of various momenta in the longitudinal direction (that is, along the collision axis). Hence the longitudinal momenta of the secondary particles are not directly related to any specific plasma bubble. The average transverse momenta $\bar{p}_T$ can be so related, however, and thus provide a measure of the initial temperature: the higher p_T , the greater the momentum boost due to the plasma expansion must have been, and hence the higher the initial temperature T . Thus, using $\bar{p}_T$ as a measure of temperature and the multiplicity $\bar{N}$ in a given momentum range as a measure of energy density, from Fig. 2, the $\bar{N}$ dependence of $\bar{p}_T$ should be as shown in Fig. 6¹⁹. Indeed, there are first indications for such a behaviour, obtained from the study of nuclear collisions in data from cosmic rays²⁰.

For a survey of experimentally observable signatures see ref. 21. As this article was being completed, a further and apparently rather clear signal for plasma formation was developed²². If deconfinement occurs, then the production of J/Ψ mesons in nuclear collisions must be very strongly suppressed, because within the plasma the charmed quark constituents of the J/Ψ cannot form a bound state, and the large mass of the J/Ψ prevents it from being formed in a later stage. No other mechanism for such a suppression is yet known. It should be emphasized, however, that with the study of dense hadronic systems we enter an unexplored region of physics, and thus surprises are to be expected.

The experimental study of high-energy nuclear collisions is scheduled to start this year. At Brookhaven National Laboratory a beam of ^{32}S ions at 15 GeV per nucleon should become available in October 1986, and three major project groups plan to perform experiments with this beam. At CERN a beam of ^{16}O (perhaps also ^{32}S) ions, of energies up to 200 GeV per nucleon, will be used by five major experimental groups in a three-week run starting in November 1986. In both laboratories there will be additional smaller-scale emulsion experiments. It is certainly hoped that these studies may provide first indications of the quark plasma; nevertheless in both cases the ions are not very heavy, and A should be large to obtain as 'macroscopic' a system as possible. At Brookhaven a booster facility is under construction, which will allow the acceleration of ions as dense as gold (^{197}Au); it should go into operation in around 1989. Beyond these projects, which rely on existing accelerators both at BNL and CERN, there are plans for the construction of a machine dedicated to high-density hadron physics at Brookhaven. This relativistic heavy ion collider (RHIC) would provide

a colliding nuclear beam facility, with energies of around 100 GeV per nucleon in each beam for even the heaviest ions. A first decision on this proposal is expected soon.

Why do a rapidly growing number of physicists find these prospects so exciting? It appears to be within reach to study

the states of matter in strong interacting physics, and in this study to recreate experimentally in the laboratory a crucial stage in the development of the early Universe.

This work was supported by the US Department of Energy contract DE-ACO2-76CH00016.

1. Mott, N. F. *Proc. phys. Soc.* **A62**, 416-422 (1949).
2. Mott, N. F. *Rev. mod. Phys.* **40**, 677-683 (1968).
3. Fritzsche, H., Gell-Mann, M. & Leutwyler, H. *Phys. Lett.* **47B**, 365-368 (1973).
4. Weinberg, S. *Phys. Rev. Lett.* **31**, 494-497 (1973).
5. Gross, D. J. & Wilczek, F. *Phys. Rev.* **D8**, 3633-3652 (1973).
6. Yang, C. N. & Mills, R. L. *Phys. Rev.* **96**, 191-195 (1954).
7. Wilson, K. *Phys. Rev.* **D10**, 2445-2459 (1974).
8. Creutz, M. *Phys. Rev.* **D21**, 2308-2315 (1980).
9. Satz, H. A. *Rev. nucl. Part. Sci.* **35**, 245-270 (1985).
10. K. Binder (ed.) *Monte Carlo Methods in Statistical Physics* (Springer, Berlin, 1979).
11. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. *J. chem. Phys.* **21**, 1087-1092 (1953).
12. Çelik, T., Engels, J. & Satz, H. *Nucl. Phys.* **B256**, 670-686 (1985).

13. Kogut, J., Polonyi, J., Wyld, H. W., Shigemitsu, J. & Sinclair, D. K. *Phys. Rev. Lett.* **53**, 644-647 (1984).
14. Engels, J. & Satz, H. *Phys. Lett.* **159B**, 151-154 (1985).
15. Berg, B., Engels, J., Kehl, E., Walti, B. & Satz, H. *Z. Phys.* **C31**, 167-174 (1986).
16. Kajantie, K. (ed.) *Quark Matter 1984, Lecture Notes in Physics* no. 221 (Springer, Berlin, 1985).
17. Bjorken, J. D. *Phys. Rev.* **D27**, 140-151 (1983).
18. Baym, G. *Quark Matter 1984, Lecture Notes in Physics* no. 221 (ed. Kajantie, K.) 39-52 (Springer, Berlin, 1985).
19. Van Hove, L. *Phys. Lett.* **118B**, 138-140 (1982).
20. Miyamura, O. *Quark Matter 1984, Lecture Notes in Physics* no. 221 (ed. Kajantie, K.) 187-195 (Springer, Berlin, 1985).
21. Gyulassy, M. *Nucl. Phys.* **A418**, 59c-85c (1984).
22. Matsui, T. & Satz, H. *Phys. Lett.* **178**, 416-422 (1986).

ARTICLES

Organization of anterior pattern in the *Drosophila* embryo by the maternal gene *bicoid*

Hans Georg Frohnhöfer & Christiane Nüsslein-Volhard

Max Planck Institut für Entwicklungsbiologie, Spemannstrasse 35, 74 Tübingen, FRG

Embryos derived from bicoid (bcd) females lack head and thorax. Cytoplasmic transplantation experiments reveal that bcd⁺-dependent activity is localized at the anterior pole of wild-type embryos. The activity can induce anterior development in mutant embryos at any position along the antero-posterior axis and suppresses posterior development.

WHEN the young *Drosophila* embryo proceeds through the stages of syncytial blastoderm towards gastrulation, the perivitelline layer of somatic nuclei looks almost homogeneous, with no morphological signs of cellular specification. But when the blastodermal layer is examined with molecular probes, the spatially restricted transcription of segmentation genes reveals that it is already being subdivided into segmental units¹⁻³. It has frequently been proposed that maternally encoded functions provide positional information in the egg and control the spatial pattern of zygotic gene expression, which in turn determines the developmental fate of individual cells. In insects, there is evidence for two centres of activity, located at the egg poles, which organize the antero-posterior axis (reviewed in ref. 4). In *Euscelis* transposition of posterior polar plasm within the egg can induce a duplicated posterior end with inversed polarity⁴. An analogous activity was found at the anterior in chironomid midges; manipulation of the anterior pole can cause double abdomen formation (reviewed in ref. 5). The activity appears to be required for anterior and to inhibit posterior development. The anterior activity in *Smittia* is sensitive to RNase and ultraviolet irradiation⁵ and are therefore characterized as RNA-containing particles.

Evidence for terminal organizing centres has been obtained in *Drosophila* by experimental removal and/or transplantation of polar cytoplasm⁶. Specific, distinct phenotypes are produced which resemble the phenotypes of several maternal effect mutants. This suggests that the genes involved are important in the establishment of the morphogenetic centres. Five loci are known in which mutations eliminate the posterior pole plasm, including the polar granules. No pole cells are formed and the development of the abdomen is abolished. Strikingly, the posteriormost somatic region of the embryo, the telson, is not

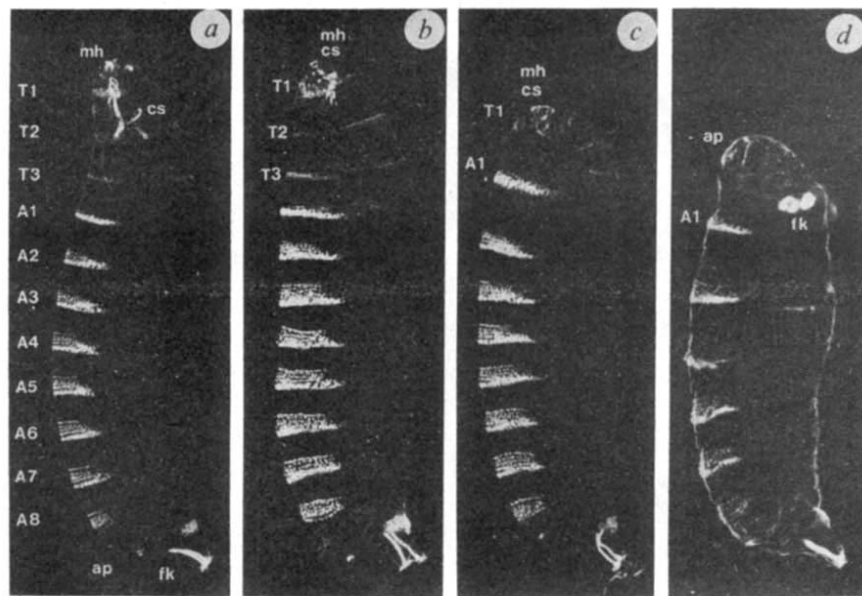
affected⁷⁻⁹. In *oskar* mutants (*osk*⁻), abdominal segmentation can be restored by transplantation of wild type pole plasm into the prospective abdominal region⁹. Only pole plasm from *osk*⁺ embryos can induce posterior development at the anterior, suggesting that this group of genes is involved in the organization of posterior pattern^{6,9}.

The anterior structures, head and thorax, are eliminated by maternal effect mutations of the bicoid group¹⁰⁻¹². Head and thorax are replaced by a second duplicated abdomen (including telson) of reversed polarity. Bicoid embryos are produced by mutations in a variety of complementation groups with quite different genetic properties, but in no case with 100% penetrance¹². Therefore these mutants do not readily define functions specifically required for anterior development. Here we describe the maternal effect gene *bicoid* (*bcd*). Embryos produced by mutant females lack head and thorax and have a second duplicated telson at their anterior end, but usually no abdominal segments of inversed polarity. We show that the function of *bcd* is mediated by a cytoplasmic activity, localized at the anterior, with long range organizing effects on the antero-posterior pattern.

Phenotype of *bicoid*

All eleven *bcd* alleles are strictly maternal. The penetrance is 100%, but the alleles differ in the extent of anterior pattern defects they produce (Fig. 1). Embryos from strong mutant alleles completely lack the head and thorax (Fig. 1d); instead they have a second telson at the anterior end. Deletions and fusions of segments are also found in the anterior abdominal region. Occasionally abdominal segments with inversed polarity are duplicated at the anterior, together with the telson. The weakest alleles show pattern defects only in the anterior head

Fig. 1 The phenotype of *bicoid*. Cuticular preparations of embryos of different mutant strength. *a*, Wild type. The weakest *bcd* alleles (*bcd*²⁻¹³, *bcd*^{E5}, and *bcd*¹¹¹) lack only labral derivatives (median tooth, dorsal bridge) and are not shown. *b*, Weak intermediate phenotype (*bcd*^{E3}/*bcd*^{E3}, 18°). The entire head is reduced; thoracic and abdominal segmentation is normal. *c*, Strong intermediate phenotype (*bcd*^{E3}/*bcd*^{E3}, 25°). The head is further reduced and the thorax is affected. On the ventral side denticle belts T2 and T3 are deleted and T1 is reduced in size. Three dorsal bulges with segmental arrays of dorsal hairs demonstrate that the thoracic region is still composed of all three segments fused together. Intermediate alleles: *bcd*^{E3} (strongly temperature-sensitive), *bcd*^{E4} and *bcd*⁰⁸⁵. *d*, Strong *bcd* phenotype (*bcd*^{E1}/*bcd*^{E1}). Head and thorax are lacking altogether. They are replaced by a complete telson with anal plates, tuft, spiracles and filzkörper, but these are not stretched. The anterior abdomen shows segment defects and fusions. Strong alleles: *bcd*^{E1}, *bcd*^{E2}, *bcd*³³⁻⁵, *bcd*²³⁻¹⁶, *bcd*^{GB}. A more detailed description of the individual alleles and their temperature dependence will be given elsewhere. Mutant alleles: *Bicoid* alleles were induced with ethyl-methanesulfonate in a marked chromosome and isolated: (1) in a large scale mutagenesis experiment designed for recessive female sterile mutations on the third chromosome (C.N.-V., G. Jürgens, K. Anderson, R. Lehmann & H.G.F., unpublished, see also refs 9, 24) (*bcd*^{E1}-*bcd*^{E5}, *bcd*⁰⁸⁵); (2) as alleles which failed to complement *bcd*^{E3} (*bcd*²⁻¹³, *bcd*¹¹¹, *bcd*³³⁻⁵, *bcd*^{GB}) or (3) from individual F₁ females heterozygous for a mutagenized chromosome, on the basis of the dominant haploinsufficient phenotype (anterior shift of the head fold) (*bcd*²³⁻¹⁶, *Df*(3R)*LIN*). The *bcd* gene is not complemented by *Df*(3R)*Scr*, *Df*(3R)*LIN*, or *Df*(3R)9A99, but covered by *Df*(3R)*Ns*^{R+17} and *Df*(3R)4SCB which places it in 84A into the same cytogenetic interval which contains *zerknüllt*²⁵⁻²⁷. **Methods.** Embryos were collected on yeast agar plates, dechorionated and devitelinated. They were mounted in Hoyer's mounting medium, diluted 1:1 with lactic acid and cleared by overnight incubation at 60 °C. Darkfield photographs. A1-A8 = ventral abdominal denticle bands, ap = anal plates, cs = cephalopharyngeal skeleton, fk = filzkörper, mh = mouth hook, T1-T3 = thoracic denticle bands. Anterior at the top, ventral at left. For the morphology of the larval cuticle see refs 28, 29.



region. With increasing allelic strength the entire head region (Fig. 1*b*), then the thorax (Fig. 1*c*) are affected. Otherwise there is no clear order in which structures are eliminated. Three anterior regions with different sensitivity for *bcd*⁺ reduction can be distinguished; the preantennal head region, antennal and gnathocephalic head segments, and thorax.

In transheterozygous combinations, *bcd* deficiencies behave like strong alleles. The allelic series thus reflects a continuous decrease of *bcd*⁺ activity (hypomorphic states); the strong alleles approximate the complete loss of *bcd*⁺ function (amorphic condition).

Development

Bicoid embryos develop normally until blastoderm formation. Morphological alterations in mutant embryos start at gastrulation. The cephalic fold, indicating the position of the gnathal anlagen, is a good landmark for changes of organization (Fig. 2). Females carrying a duplication or a deletion for *bcd*⁺ produce embryos with the cephalic fold shifted towards a more posterior or anterior position respectively (Fig. 2*a-c*). With further decrease of *bcd*⁺ activity in homozygous alleles, the cephalic fold becomes weaker, and finally disappears (Fig. 2*d*). At intermediate mutant strength a posterior midgut plate is formed at the anterior end, indicating a duplicated telson. This is seen more clearly in strong alleles (Fig. 2*e*). However, no pole cells are formed at the anterior end and no pole plasm is found there in the early stages. The shifted cephalic fold reveals that the *bcd* phenotype is associated with changes in the blastodermal anlagen plan. Thus the reduction in the size of anterior structures in the differentiated embryo can be traced back to shrinkage of the corresponding primordia. The abdominal anlage is enlarged and spreads towards the anterior (data not shown).

Localization of *bcd*⁺ activity

We tried to restore the *bcd* function in strong mutant embryos by transplantation of cytoplasm from wild type embryos. Cytoplasm was taken from the anterior tip of young (stage 2)¹³ wild type embryos and transplanted into the anterior of mutant embryos of the same age. The amount of transplanted material corresponded to about 5% of the egg volume; one wild type embryo was used as a donor for 1-2 mutant recipient embryos. Nearly all recipient embryos showed a change in phenotype from strong *bcd* towards normal. Some embryos formed only thoracic segments (Fig. 3*b*); others, in addition to thorax, gained almost complete heads (Fig. 3*a*). Three exceptional embryos survived until pupation. The anterior abdomen, which is misshapen in the mutant, became normalized whenever thorax was induced. However, the normal number of abdominal segments was often not restored; new defects arose in the midabdominal region. The embryos were classified according to the degree of anterior development observed (Fig. 4).

The rescuing activity of anterior cytoplasm is strongly dependent on the genotype of the female producing the donor eggs (Fig. 4*b*). Cytoplasm from *bcd*⁻ embryos has no rescuing activity. Eggs laid by females with only one *bcd*⁺ gene (*Df bcd*⁻/*bcd*⁺), the normal two doses (*bcd*⁺/*bcd*⁺) and three doses of the gene (*Dp bcd*⁺/*bcd*⁺/*bcd*⁺) have cytoplasm with increasing *bcd*⁺ activity. Varying the amount of transferred cytoplasm mimicks the effect seen with varying *bcd*⁺ gene dosage (Fig. 4*a*), suggesting a linear relationship between *bcd*⁺ gene expression and transplantable activity. This suggests that the *bcd*⁺ gene product might be the active transplanted factor and that no negative feedback acts on its accumulation at the anterior end.

The efficiency of phenotypic rescue decreases with increasing age of both donors and recipients (Fig. 4*c, d*). Unfertilized eggs

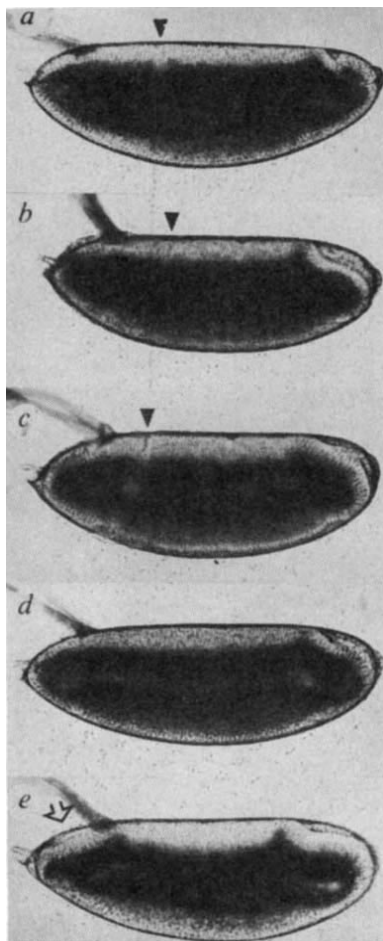


Fig. 2 Gastrulation pattern of mutant embryos. Photographs of living embryos produced by females with: *a*, three *bcd*⁺ copies (*C(1)RM.y; T(Y;3)MA9^P/T(Y;3)A109^D*)^{30,31}; *b*, two *bcd*⁺ copies (wild type); *c*, one *bcd*⁺ copy (*bcd*^{GB/+}); *d*, *bcd*^{E4}/*bcd*^{E4}, intermediate allele; *e*, *bcd*^{E1}/*bcd*^{E1}, a strong allele. The position of the cephalic furrow (arrowhead) varies with *bcd*⁺ activity between 62% EL (*a*), 65% (*b*) and 71% (*c*). The significance of the backshift in *a* is obscure because females carrying a duplication tend to produce eggs with an unusual shape. Both types of embryos with shifted cephalic folds usually produce normal patterns and survive to adulthood. The cephalic fold is absent in intermediate and strong alleles (*d*, *e*). In the strong phenotype (*e*), a second posterior midgut plate is evident at the antero-dorsal side of the embryo (open arrow). Later in development it performs characteristic invagination movements. Embryos were collected on agar plates, covered with Voltalef 3S oil and photographed in transmitted light with a Zeiss photomicroscope. Anterior left, dorsal at the top.

contain the most active cytoplasm. The activity in older embryos decreases slowly until the last cleavage, after which the activity is sharply reduced (Fig. 4c). The high activity observed in unfertilized eggs persists for much longer (data not shown). The responsiveness of recipients decreases continuously (Fig. 4d), suggesting that the *bcd*⁺ activity is involved in a process which lasts from fertilization to the end of cleavage.

To determine the localization of the *bcd*-dependent activity in wild-type embryos, cytoplasm was taken from various positions each covering about 10% egg length (EL) along the antero-posterior axis. Figure 4e shows that *bcd*⁺ activity is high at the anterior pole and declines sharply away from it, suggesting a graded distribution of *bcd*⁺ activity in the anterior third of the egg. However, the resolution of 10% EL is at the limit of our technique, and a high homogeneous distribution from 85–100% EL would also fit the data. No dorso-ventral difference in activity and no clear difference between central yolk and cortical cytoplasm was obvious. In older donors (syncytial blastoderm), the prevalent anterior localisation is still present (data not shown).

Induction of anterior development

Anterior structures can be induced by injecting *bcd*⁺ activity at any position along the antero-posterior axis. The anterior structures are formed near the injection site. The corresponding reorganization of the pattern involves major changes in embryonic polarity. The abdomen also reorganizes according to the conditions installed by the anterior determinants. The ability of the injected material to induce new thoracic or cephalic structures strongly decreases with distance from the anterior pole (Fig. 5f). The posterior pole seems to exert a suppressive force which declines toward the anterior end.

We transferred increased amounts of cytoplasm to compensate for reduced efficiency in induction in the posterior half of

the egg. The results are shown together with those of standard experiments in Fig. 5a–f. We now describe some typical examples of the phenotypes created by heterotopic transplantation, starting with the phenotypic response upon anterior injection.

Even embryos rescued by anterior injection display slight deviations from the phenotypes of the allelic series. They tend to have a rather poorly developed head compared to the thorax (Fig. 3b). It may be that when the material is transplanted it is not deposited in the very anterior tip but is spread over an area around 85% EL. If so, the distribution of the activity influences the relative sizes of head and thorax. Injections at 67% EL produced about the same range of phenotypes as anterior injection.

Injection at 50% EL induces head structures near the middle of the injected embryo, flanked by two thoracic regions with opposite polarity (Fig. 3c, e, g). Telson structures characteristic of the mutant phenotype are still present at both poles. A maximum of anterior positional information is thus established near the site of injection in the centre of the egg and decreases in both directions toward the poles. Abdominal segments are only formed at the original posterior site and are frequently reduced in number. The phenotypes suggest that the blastodermal region giving rise to abdomen has been compressed into a very posterior region of the egg.

The asymmetry observed in the embryos is also seen in the inverted region of the induced thorax, which tends to have fewer segments than the one with normal polarity. In the borderline case there is no inverted region in the induced thorax, suggesting a general preference to establish anterior maxima at the terminal region of the egg rather than in the middle. Further, the suppression by posterior factors may primarily affect the posterior portion of the injected material and cause anterior maxima to be established at blastodermal positions somewhat anterior to the site of injection.

Injectations to more posterior positions (15% EL and 32% EL) can induce head structures together with three thoracic segments at the posterior end of the egg, the polarity of the entire embryo being reversed. At the same time formation of abdominal segments is strongly suppressed, never leaving more than one field of abdominal denticles. Complete suppression of the abdomen is observed more frequently after injection at 32% EL than at 15% EL (Fig. 5a-f), suggesting that *bcd*⁺ activity probably does not act on the posterior pole plasm to suppress abdomen specific factors⁹. Rather it appears to act on a signal spreading from the pole or to share common targets with this signal.

When abdominal segments are formed in phenotypically less extreme animals, the posterior segments have inversed polarity. Figure 3d, f shows an embryo in which, in addition to producing thoracic structures, two inversed abdominal segments were formed, the posterior one having the appearance of a first abdominal segment. As the anterior activity cannot induce abdominal development on its own, it must have modified the action of some abdomen specific factors. Such an interaction would ensure a continuous change of morphological values at the boundary of thorax and abdomen.

Phenocopies by cytoplasm removal

The *bcd* phenotype can also be produced (phenocopied) by elimination of anterior cytoplasm. Pricking the anterior end of

wild type eggs (stage 2) to allow leakage of ~5% volume mainly generates embryos with reduced head structures, resembling weak *bcd* phenotypes (Fig. 6a). Treating embryos in which *bcd*⁺ activity is already reduced causes stronger *bcd* phenocopies approaching the amorphic phenotype (Fig. 6b, c). Strong mutant embryos do not show any change on loss of anterior cytoplasm (Fig. 6d). The artificial phenotypes also show abdominal defects suggesting that the entire *bcd* phenotype is caused by the lack of anteriorly localized cytoplasmic activity. Pricking preblastodermal eggs at non-terminal sites does not eliminate cuticular pattern elements, whereas leakage at the posterior pole with removal of the posterior pole plasm causes abdominal defects resembling weak *oskar* phenotypes⁶. In *Smittia*, in contrast to *Drosophila*, anterior pricking creates bicaudal embryos with a reversed abdomen at the anterior¹⁴. This difference reflects a general tendency in *Smittia* to produce symmetrical states, if one terminal activity is not balanced by its antagonist.

Conclusions

Mutant *bcd* embryos fail to produce head and thorax. Corresponding primordia do not appear in the anlagen plan at the blastoderm stage. Instead a second telson is found at the anterior end and the abdominal primordium is expanded towards the anterior. The *bcd* phenotype is thus composed of two types of defect; loss of anterior structures due to shrinkage of the anlagen

Fig. 3 Phenotypes of *bcd*⁻ embryos injected with anterior cytoplasm. *a, b*, Strong and weak response to anterior injection. The embryo in *a* has gained all anterior structures with the exception of the labral derivatives (median tooth and dorsal bridge). Note that the anterior abdominal region is also normalized; but mid-abdominal segments are fused, a type of defect not seen in non-injected controls. The embryo in *b* developed two thoracic segments, but no head structures; small anal plates are retained at the anterior. In the phenotypic series, embryos with more than one thoracic segment also have some remnants of head structures (compare Fig. 1c). Moreover injected embryos display T2 and T3, but the denticle belt of T1 is the most stable to mutation. These discrepancies are interpreted as being caused by a subterminal and less compact positioning of the activity in the injected animals. *c, e, g*, Embryo after injection at 50% EL. The induced cephalic structures are flanked at the anterior and at the posterior side by thoracic regions of opposite polarity. Most of the head structures are formed at the dorsal side of the embryo (shown in detail in *g*). Mainly maxillary derivatives can be recognized. In this type of embryo, the head is not oriented along the antero-posterior axis of the egg; cephalic structures frequently have dorsal or lateral positions and their sequence seems to be delineated tangentially along the circumference of the egg within a transverse plane. The transverse plane is the plane of bilateral symmetry seen among the head structures. Sometimes pairs of head structures occur at either side of the embryo, (4 mouth hooks, 4 maxillary sense organs and/or 4 groups of ventral cirri in total). Due to cuticular holes and the scarcity of characteristic structures other than those of the maxilla the anteroposterior polarity of the head is difficult to assess. Possibly the anterior end of the head lies either dorsally or ventrally depending on the localization of the injected cytoplasm. *d, f*, Embryo injected at 15% EL. Two bands of thoracic denticles have been induced at the posterior. They are the anteriormost elements of an inversed region including about 2.5 abdominal segments. Anteriorly the original abdominal polarity is preserved. The appearance of a new A1 shows that not only the polarity but also the quality of abdominal segments has been changed. The posterior telson is partially suppressed. Arrows, segment polarity; polarity is ambiguous in the case of the broken arrow in *e*. Black dots and triangles, thoracic sensory organs. Identification of labial (li) and antennal (an) structures is not certain in the specimen in *e*. cir, cirri; mso, maxillary sense organ. Other designations as for Fig. 1.

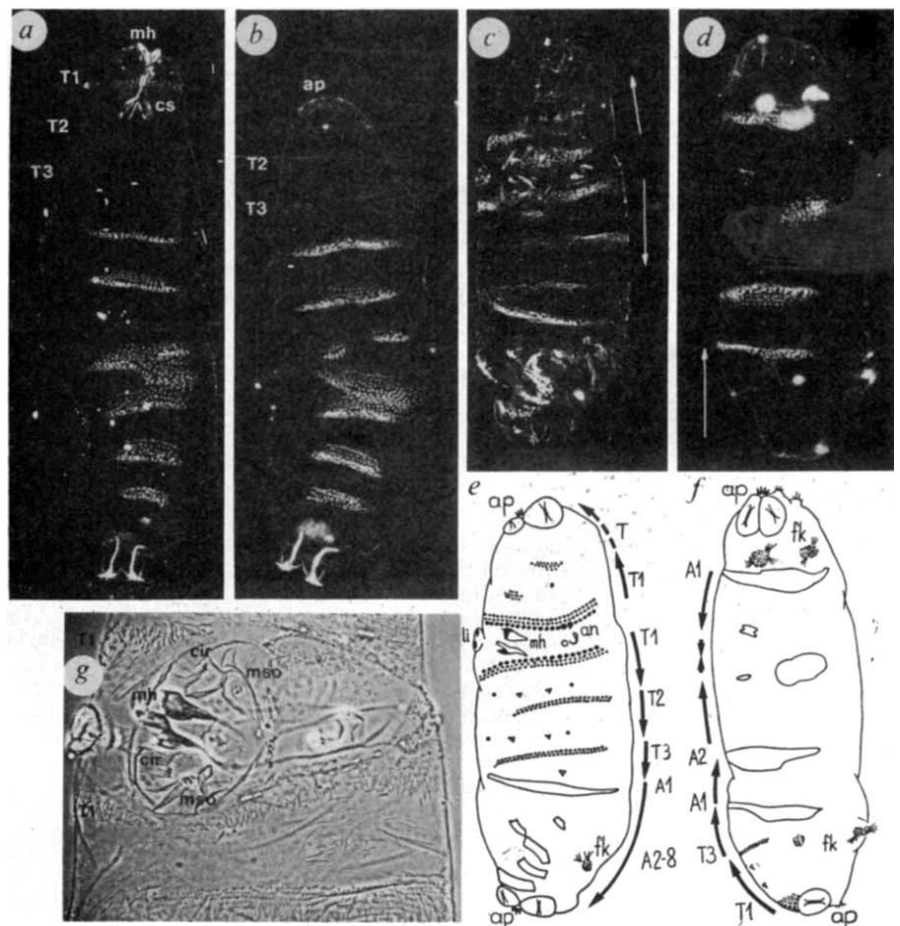
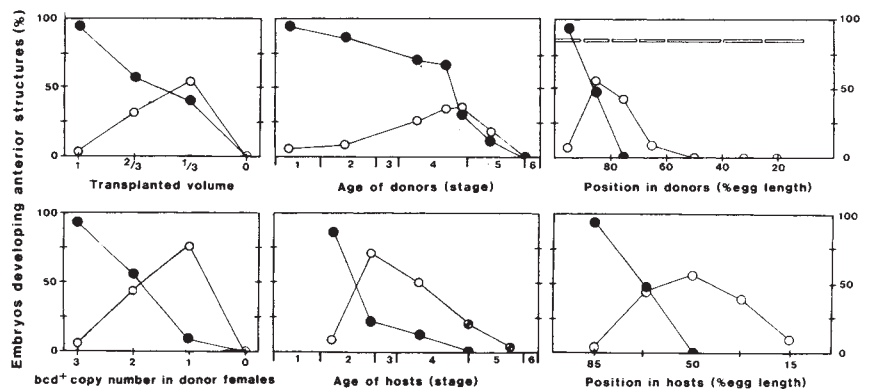


Fig. 4 Phenotypic response on cytoplasmic injection. Frequencies of embryos induced to develop thoracic and cephalic structures (●, strong response) or displaying thoracic structures only (○, weak response) provide a measure for efficiency of induction obtained under various experimental conditions. Females carrying a duplication for *bcd*⁺ produced the donor eggs and the recipients were from *bcd*^{E1}/*bcd*^{E2} females unless otherwise indicated. Donor eggs were of stage 1 or 2 in *d* and unfertilized elsewhere; recipient eggs were of stage 1 or 2. Cytoplasm was withdrawn from an anterior area (80–100% EL). When injected into an anterior position it was observed to spread in the central core of the egg between 75 and 95% egg length. The amount of cytoplasm transferred was about 5% of the egg volume. Average number of embryos for each point was 33 (range 13–54). *a*, The usual amount of injected cytoplasm is set to 1 arbitrary unit. The different relative volumes (1, 2/3, 1/3) were achieved by counting the time cytoplasm was allowed to flow from the needle into the egg (15 s, 10 s, 5 s respectively). *b*, Donor females were of the genotypes *bcd*⁺/*bcd*⁺/*bcd*⁺ (*T(Y:3)MA9^P/T(Y:3)A109^D*), wild type, *bcd*⁺/*bcd*⁺ (*Df(3R)9A99/+*), and *bcd*^{E1}/*bcd*^{E2} with *bcd*⁺ copy number decreasing from three to zero. *c*, *d*, Stages as described in methods. Unfertilized eggs are taken to represent the youngest stage 1 in *c*. In old recipients (stage 5) despite the low frequency of induction affected animals mostly display some remnants of head structure surrounded by few thoracic denticles (●). *e*, Bars on top represent the areas from where cytoplasm was withdrawn. *f*, Further information about the phenotypes obtained in this experiment is provided in Fig. 5*a–c*.

Methods. Eggs were dechorionated with NaOCl and attached to a sticky cover slip. Recipient embryos were dried for 10–12 min with a fan, both donors and recipient eggs were then covered with halocarbon oil (Voletaf 10S). Injections were performed under a compound microscope. Micropipettes of about 12 µm external diameter driven by a syringe were used. Recipient eggs were entered by the micropipette at the anterior pole in most cases, sometimes from the side but never from posterior in order to avoid loss of the posterior pole plasm. To assess the amount of cytoplasm transferred, the entire content of an egg was drawn into the needle. About 20 recipient embryos could be injected using one donor embryo. Cytoplasm which leaked out during injection was removed afterwards. Experiments were performed at 18 °C or 20 °C. Usually about 80% of injected embryos differentiated cuticle and could be scored. The stages are those of Campos-Ortega and Hartenstein¹³ (1, freshly laid egg; 2, early cleavage; 3, pole cell formation; 4, syncytial blastoderm (nuclear cycle 10–13)³²; 5, blastoderm cellularization; 6, gastrulation). All embryos were staged visually during injection; young and old stage 2 embryos in *d* were distinguished by the time they entered gastrulation.

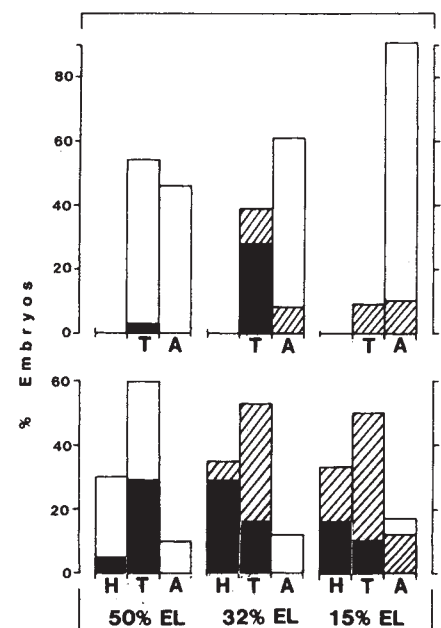


and abdominal segmentation defects in a stretched posterior domain. The action of *bcd*⁺ on both the anterior and abdominal region is mediated by the anteriorly localized activity. The anterior centre thus influences the pattern of more than half of the egg length in spite of its striking anterior localization. The structures of the telson occupy the terminal positions in the egg if there is low *bcd*⁺ activity, irrespective of which other structures become their neighbours.

The amount of *bcd*⁺ activity present in the egg determines the size of cephalic and thoracic anlagen. It also affects the degree of anteriority in the cuticular structures. The exact positioning of the *bcd*⁺ dependent activity seems to influence

the relative preponderance of thoracic and cephalic elements in homotopic transplantations. Changes of polarity associated with heterotopic transplantations provide more striking evidence that the *bcd*⁺ activity not only establishes the anterior primordium but also controls its internal organization. Describing the action of *bcd* therefore requires more than one spatial parameter. The production of different anterior qualities depending on quantity and distribution of *bcd*⁺ activity is best explained by assuming a graded distribution of factors determining the positions of head and thorax. The well organized polarity observed in the cuticle of injected embryos contrasts with the coarse conglomeration of activity present immediately after transplantation. The

Fig. 5 The effect of heterotopic transplantation on abdomen formation. The standard volume (*a, b, c*) or 1.5 × standard volume (*d, e, f*) of anterior cytoplasm was transplanted into three different positions at 50% EL (*a, d*), 32% EL (*b, e*) and 15% EL (*c, f*) of *bcd*^{E1}/*bcd*^{E2} embryos; other conditions were as described for Fig. 5. On an average 35 embryos have been recorded for each panel. Columns in each panel show the frequency of embryos with head and thoracic structures (H), only thorax (T), or neither of the two (A, abdomen). The marked areas in each column indicate the behaviour of the abdomen occurring in each class. Black portion, embryos with complete abdominal suppression (without abdominal denticle belts); hatching, cases in which inversed abdominal segments have been induced; white, the remainder of the embryos with an abdomen (partially suppressed or not). The phenotype expressed may depend on whether the transferred cytoplasm happens to be closely centred around the injection site or spreads some way through the egg. A certain amount of cytoplasm may be sufficient to induce head if centred well at 50% EL, but will then not suppress the abdomen very strongly. If the same amount of cytoplasm induced only thorax, abdominal suppression was much stronger suggesting less restricted localization (*d*). In many of the embryos displaying thorax only, segmental organization was poor and polarity was disordered. Otherwise we observe the expected positive correlation between transplanted volume, extent of anterior inductions, and suppression of abdomen. Abdominal suppression is most effective at 32% EL, whereas abdominal inversion is obtained with higher frequencies at the posteriormost injection site. The differences between the three sites tend to decrease with higher amounts of transplanted cytoplasm, probably due to lower spatial resolution. Once the threshold of posterior inhibition is overcome, anterior induction may be favoured in terminal positions. On transferring large volumes of cytoplasm, frequencies of polarity reversal within induced anterior regions varied with the site of injection thus: 2% (85% EL), 12% (67% EL), 55% (50% EL), 6% (32% EL), 5% (15% EL).



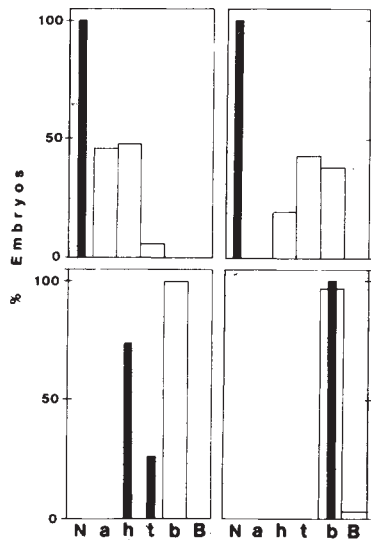


Fig. 6 *a-d*, Production of *bcd* phenocopies by removal of cytoplasm. The phenotypic response upon pricking and subsequent leakage of anterior cytoplasm from different genotypes is shown. Maternal genotype of pricked eggs: *a*, wild type; *b*, *Df(3R)9A99/+; c, bcd^{E3}/bcd^{E3}*; *d, bcd^{E1}/bcd^{E2}*. The letters along the abscissa designate the phenotypic classification. N, essentially normal; a, head weakly reduced with antenna still present; h, head strongly reduced; t, thorax reduced (head reduced or absent, usually some telson structures duplicated); b, head and thorax entirely absent, mimicking the *bcd*⁻ phenotype; B, some abdominal segments inversed reminiscent to *Bicaudal* (only 1 case). Black columns, phenotypic spectrum of untreated embryos; white columns pricked embryos. Removal of anterior cytoplasm causes a shift toward the amorphic phenotype. Treating amorphic mutant embryos elicits no further change (for example, toward *bicaudal* embryos).

Methods. Embryos were mounted as in transplantation experiments, but not dried. Anterior pricking caused loss of cytoplasm (5% egg volume), which was carefully removed from the outside of the eggs. About 70% of the embryos survived to the differentiated stage⁶. At least 30 embryos have been recorded for each panel.

transplanted activity may achieve some redistribution, but we feel that it probably acts as a localized source providing a spreading signal to instruct the peripheral sheet of nuclei. This might explain why the position of the transplanted material is crucial for the eventual phenotype but is not necessarily isomorphic with the pattern induced. An anteriorly localized RNA transcribed from the *bcd* locus and distributed similarly to our transplantable activity has been found¹⁵. The translation product may be the postulated signal.

The spatially restricted expression of zygotic segmentation genes of the 'gap' class might be the focus of *bcd* function in the anterior half of the embryo^{3,16}. Gap mutants eliminate continuous blocks of segments (or parasegments¹⁷) along the body axis¹⁸⁻²⁰. In contrast to the *bcd* phenotype, the anlagen plan outside the regions deleted suffers no major alterations²⁰⁻²². In *hunchback* mutants, labium and thorax (the posterior part of the *bcd* domain) are lacking²⁰. Therefore there may be no expression of *hunchback* in *bcd* embryos. Mutants in *Krüppel* (*Kr*) eliminate segments T1-A5, segments which are dependent on the function of *bcd*, as well as segments dependent on *oskar*. *Kr* should therefore be expressed in both *bcd* and *oskar* embryos. However, the abdominal expansion seen in *bcd* embryos suggests an anterior shift in the spatial expression of *Kr*.

Abdominal development is strictly dependent on maternal genetic functions as exemplified by the gene *oskar*, which has been examined by means quite similar to our own work on *bcd*⁹. In controlling the posterior pole plasm and abdominal development, *oskar* and its relatives^{7,8} constitute a posterior analogue to the function of *bcd*. However the interplay between *bcd* and *oskar* dependent factors is not symmetric. Anterior anlagen show no major expansion towards the posterior in *oskar* embryos^{9,33}, so the establishment of the anterior *bcd* domain is independent

of the posterior centre. But an antagonistic effect can be provoked experimentally; when posterior pole plasm is transplanted to the anterior tip of wild-type eggs it suppresses anterior development^{6,9}. Anterior cytoplasm transplanted into the range of the future abdomen is weakened in its inductive strength (Fig. 4f). Because the anterior source material seems to be outside the range of inhibiting posterior factors in wild-type eggs, there is no obvious natural function for such an interaction during postfertilization stages, when both polar centres are already established.

The *bcd*⁺ gene is not required to establish abdominal anlagen as such but is required to confine abdominal development to the posterior half of the embryo. Depending on position, *bcd*⁺ activity can drive abdominal primordia backwards in the blastodermal anlagen plan; change abdominal segmental qualities toward more anterior values; or suppress abdominal development. Thus *bcd*⁺ exerts profound interventions on structures which develop fairly autonomously in *bcd*⁻ embryos. If one assumes abdominal development to be organized by a posterior gradient emanating from the pole plasm²³, then inhibition and modification of this gradient by *bcd* dependent activity would explain all types of interactions described here. Alternatively, *bcd* might not act on maternally encoded factors and contribute nothing to abdominal development other than to determine its anterior boundary by control of *Kr* gene expression.

We thank Ruth Lehmann for discussions and Kavita Arora, Thomas Berleth, Stephen Cohen, Martin Klingler, David Marcey and Siegfried Roth for helpful suggestions on the manuscript. Ruth Lehmann, Kathryn Anderson and Gerd Jürgens participated in the mutant screens that led to the isolation of the *bicoid* alleles. We thank Gunilla Weinraub for drawings, Roswitha Groemke-Lutz for the photographic prints, and Dominique Kaiser for typing the manuscript.

Received 13 August; accepted 8 October 1986.

- Hafen, E., Kuroiwa, A. & Gehring, W. J. *Cell* **37**, 833-841 (1984).
- Ingham, P. W., Howard, K. R. & Ish-Horowitz, U. *Nature* **318**, 439-445 (1985).
- Knipple, D. C., Seifert, E., Rosenberg, U. B., Preiss, A. & Jackle, H. *Nature* **317**, 40-44 (1985).
- Sander, K. *Adv. Insect Physiol.* **12**, 125-238 (1976).
- Kalthoff, K. in *Determinants of Spatial Organization* (eds Subtelny, S. & Konigsberg, I. R.) 97-126 (Academic, New York, 1979).
- Frohnhofer, H. G., Lehmann, R. & Nüsslein-Volhard, C. *J. Embryol. exp. Morph.* (in the press).
- Boswell, R. E. & Mahowald, A. P. *Cell* **43**, 97-104 (1985).
- Schüpbach, T. & Wieschaus, E. *Wilhelm Roux Arch. Dev. Biol.* **195**, 302-317 (1986).
- Lehmann, R. & Nüsslein-Volhard, C. *Cell* (in the press).
- Bull, A. L. *J. exp. Zool.* **161**, 221-241 (1966).
- Nüsslein-Volhard, C. *Wilhelm Roux Arch. Dev. Biol.* **183**, 249-268 (1977).
- Mohler, D. & Wieschaus, E. *Genetics* **112**, 803-822 (1986).
- Campos-Ortega, J. A. & Hartenstein, V. *The embryonic development of Drosophila melanogaster* (Springer, Heidelberg, 1985).
- Schmidt, O., Zissler, D., Sander, K. & Kalthoff, K. *Dev. Biol.* **46**, 216-221 (1975).
- Frigerio, G., Burri, M., Bopp, D., Baumgartner, S. & Noll, M. *Cell* (in the press).

- Meinhardt, H. J. *Cell Sci. Suppl.* **4**, 357-387 (1986).
- Martinez-Arias, A. & Lawrence, P. A. *Nature* **313**, 639-642 (1985).
- Nüsslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795-801 (1980).
- Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Dev. Biol.* **104**, 172-186 (1984).
- Lehmann, R. & Nüsslein-Volhard, C. *Dev. Biol.* (in the press).
- Carroll, S. B. & Scott, M. P. *Cell* **45**, 113-126 (1986).
- Ingham, P. W., Ish-Horowitz, D. & Howard, K. R. *EMBO J.* **5**, 1659-1665 (1986).
- Meinhardt, H. J. *Cell Sci.* **23**, 117-139 (1977).
- Anderson, K. V. & Nüsslein-Volhard, C. *Nature* **311**, 223-227 (1984).
- Lindsley, D. & Zimm, G. *Dros. Inf. Serv.* **62**, 1-226 (1985).
- Wakimoto, B. T., Turner, F. R. & Kaufman, T. C. *Dev. Biol.* **102**, 147-172 (1984).
- Jürgens, G., Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Wilhelm Roux Arch. Dev. Biol.* **193**, 283-295 (1984).
- Lohs-Schardin, M., Cremer, C. & Nüsslein-Volhard, C. *Dev. Biol.* **73**, 239-255 (1979).
- Jürgens, G., Lehmann, R., Schardin, M. & Nüsslein-Volhard, C. *Wilhelm Roux's Arch. Dev. Biol.* **195**, 359-377 (1986).
- Lehmann, R. dissertation, Univ. of Tübingen (1985).
- Lindsley, D. & Sandler, L. *Genetics* **71**, 157-184 (1972).
- Zalokar, M. & Erk, I. *J. microsc. Biol. Cell* **25**, 97-106 (1976).
- Carroll, S. B., Winslow, G. M., Schüpbach, T. & Scott, M. P. *Nature* **322**, 278-280 (1986).

A parallax effect due to gravitational micro-lensing

B. Grieger, R. Kayser & S. Refsdal

Hamburger Sternwarte, Gojenbergsweg 112,
D-2050 Hamburg 80, FRG

The astrophysical importance of gravitational micro-lensing due to stars in the deflecting galaxy is now well-known¹⁻⁷. Each macro-image (caused by the corresponding smoothed-out galaxy) can be split up into several micro-images with typical angular distances measured in micro-arcseconds. Here we discuss a parallax effect due to gravitational micro-lensing (star disturbances). We show that such an effect would most favourably be observed during high amplification events when a compact source (quasar) crosses a critical curve. During such an event even a relatively small displacement (≥ 0.1 AU) of the observer may lead to measurable differences in the image brightness. We discuss this parallax effect with special emphasis on the determination of (1) the size of the quasar and (2) the relative transversal velocities involved. It should be possible to observe such a parallax-effect with a relatively small telescope (10–20 cm) included in an interplanetary mission. In extreme cases, annual oscillations in the light curves of lensed quasars can be observed from the Earth, but in most cases these oscillations can be neglected.

Of particular interest among instances of gravitational micro-lensing are the high amplification events² that occur when a compact source crosses a critical curve¹. These high amplification events where $\Delta m \approx 3$ are suddenly initiated (or terminated) by the appearance (or disappearance) of two bright micro-images^{1,4,5}. It is a kind of 'gravitational eclipse' effect, allowing the formation of large spatial magnitude gradients and the occurrence of very fast changes in the observed brightness due to transverse motion (grad $m \approx 0.3$ AU⁻¹ and $\partial m / \partial t \approx 0.1$ day⁻¹ for a source radius of 10^{-4} pc and a typical transverse velocity of 500 km s⁻¹). The frequency of such high amplification events has been estimated² to be about 0.3 per year for all known gravitationally-lensed quasars. The time scale of such events is given approximately by the diameter of the source divided by the transverse velocity V_T of the source relative to the critical curve in the source plane. We note that V_T is equivalent to a relative transverse velocity (v_T) between the observer and the critical curve in the observer plane

$$v_T = -V_T \frac{D_d}{D_{dq}} (1 + z_d) \quad (1)$$

where D_d and D_{dq} are apparent size distances and z_d is the redshift of the deflector. Because the critical curves have a radius of curvature of the order of one normalized length unit² ($\sim 10^4$ AU for solar mass stars), we can approximate the critical curves by straight lines inside the solar system. We can also assume that the observed luminosity of a point source locally only depends on the distance d to the critical curve ($\propto d^{-1/2}$)^{1,4,8}.

During a high amplification event, the observed brightness changes very rapidly, and even a small displacement of the observer can bring about measurable changes. During a typical event ($\Delta m \approx 3$), a displacement of only 0.1 AU (perpendicular to the critical curve) in the plane of the observer leads to a change in brightness of up to $\delta m \sim 0.025$ for a typical lensed quasar ($D_d \approx D_{dq}$ and a source radius of 10^{-4} pc). (For the somewhat untypical lens object OSO 2237+0305 ($D_{dq}/D_d \sim 10$), $\delta m \sim 0.25$.)

This parallax effect raises the possibility of calculating v_T from observations made by three observers some distance apart (see Fig. 1). If the time lags δt_{AB} and δt_{AC} in the light curves

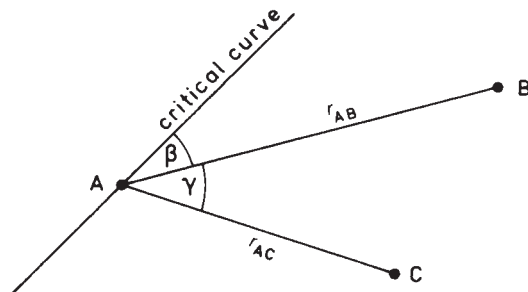


Fig. 1 Position of 3 observers, projected onto the observer plane.

are known, we can eliminate the angle β and determine the relative transverse velocity v_T (perpendicular to the critical curve) in the observer plane:

$$v_T = |\sin \gamma| \left\{ \left(\frac{\delta t_{AB}}{r_{AB}} \right)^2 + \left(\frac{\delta t_{AC}}{r_{AC}} \right)^2 - 2 \cos \gamma \frac{\delta t_{AB} \delta t_{AC}}{r_{AB} r_{AC}} \right\}^{-1/2} \quad (2)$$

Here we have made use of the fact that the time lags δt are equal to the time differences across the critical curve.

From equation (1) we obtain the corresponding velocity V_T in the source plane. From the typical time scale Δt of the high amplification event, we can then determine the time needed by the source to cross the critical curve, and thereby the radius $R = k \Delta t / V_T$ of the source, where k (≈ 1) depends on the luminosity profile of the source and the amplitude of the event. This will be discussed in a subsequent paper.

If there are only two observers, A and B, a straightforward relation between the two unknowns v_T and $\sin \beta$ emerges (see Fig. 1).

$$v_T = r_{AB} \sin \beta / \delta t_{AB} \leq r_{AB} / \delta t_{AB} \quad (3)$$

We thus see that with only two observers an upper limit of v_T is obtained.

Because reliable determination of the time lags does not require very exact observations we propose that a rather small telescope (diameter 10 to 20 cm) would be sufficient in the case of the brightest multiple quasars where $m \approx 17$. In view of the very important information that could be obtained the possibility of sending such a telescope in an interplanetary mission should be investigated.

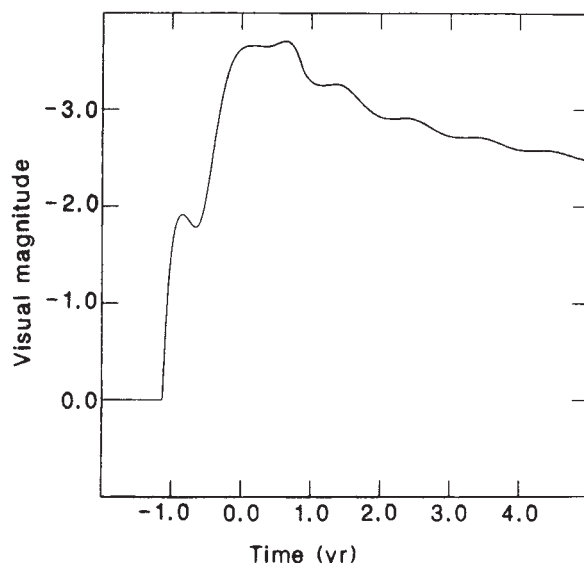


Fig. 2 Simulated light curve of a high amplification event in the lens object 2237+0305, showing oscillations due to the annual motion of the earth.

Considering the sighting of the somewhat improbable lensing by a double star^{9,10} in our galaxy ($D_{\text{dq}} \approx 10^7 D_{\text{d}}$), a parallax effect could possibly be observed by Earth-bound observers. During the most rapid variations, an observer displacement of only 6 km results in a brightness change of up to about 0.1 mag. The time scale of the high amplification event would then typically be measured in seconds.

Another possible application of the parallax effect would be to make use of the annual motion of the Earth. For an Earth-bound observer the distance $d(t)$ to the critical curve in the observer plane varies (1) due to the transverse velocity of the solar system relative to the critical curve and (2) due to the annual motion of the Earth. We note that the position of the Earth projected onto the plane of the observer is determined by the celestial coordinates of the observed object.

Figure 2 shows a simulated light curve for the lens object 2237+0305, which because of the large value of $D_{\text{dq}}/D_{\text{d}}$ (~ 10) gives the best chance of observing the effect. Here we have used a dimensionless surface density of continuously-distributed matter of $\sigma_c = 0.8$, a Hubble parameter of $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and a standard cosmological model ($q_0 = \sigma_0 = 0$). The source has a radius $R = 2.7 \times 10^{-4} \text{ pc}$ and a luminosity profile $I(r) = I_0 (1 - (r/R)^2)^{1/2}$. This seems to be more reasonable than $I(r) = \text{constant}$ and has the advantage of leading to an analytical solution of the mean amplification $\langle A(d) \rangle$. The change in brightness is given by $m(d) = -2.5 \log \langle A(d) \rangle$. We further assume a

transverse velocity of one source radius per year, that is 260 km s^{-1} in the source plane or 25 km s^{-1} in the observer plane. The annual motion of the Earth shows up clearly in the light curve as an oscillation (Fig. 2). It is clear that observable effects only occur if (1) the transverse velocity of the solar system relative to the critical curve is smaller than or about equal to the orbital velocity of the Earth ($v_T \leq 50 \text{ km s}^{-1}$) and (2) the projection $\tilde{R}$ of the source radius R onto the observer plane is not too great ($\tilde{R} \leq 100 \text{ AU}$). The first condition is particularly unlikely to be fulfilled, because a typical transverse velocity in the observer plane is of the order of 500 km s^{-1} . It therefore seems safe in most cases to neglect the annual parallax effect by gravitational micro-lensing, even during the high amplification events.

This work was supported in part by the Deutsche Forschungsgemeinschaft under Az. Re 439.3.

Received 11 June; accepted 14 August 1986.

1. Chang, K. & Refsdal, S. *Nature* **282**, 561–567 (1979).
2. Kayser, R., Refsdal, S. & Stabell, R. *Astron. Astrophys.* **166**, 36–52 (1986).
3. Gott, J. R. *Astrophys. J.* **243**, 140–146 (1981).
4. Chang, K. Ph.D thesis, Univ. Hamburg (1981).
5. Chang, K. & Refsdal, S. *Astron. Astrophys.* **132**, 168–178 (1984).
6. Young, P. *Astrophys. J.* **244**, 756–767 (1981).
7. Paczynski, B. *Astrophys. J.* **301**, 503–516 (1986).
8. Chang, K. *Astron. Astrophys.* **130**, 157–161 (1984).
9. Bartels, G. G. thesis, Univ. Hamburg (1981).
10. Schneider, P. & Weiss, A. *Astron. Astrophys.* (submitted).

Companion star of a binary millisecond pulsar

G. A. Wright & E. D. Loh

Department of Physics and Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544, USA

Binary pulsars offer important clues about the evolution of neutron stars. We have searched for and identified the companion star of the recently discovered¹ binary millisecond pulsar PSR1855+09. The optical counterpart is compatible with a $0.3M_{\odot}$ white dwarf with a surface temperature of $\sim 5,900 \text{ K}$. This is in accord with currently accepted ideas on the genesis of such a fast binary pulsar². The low temperature of the white dwarf implies that the system is $>10^9 \text{ yr}$ old. Such an old pulsar is not easily accounted for by the conventional theory, in which the pulsar ceases to be active after its magnetic field decays, a process believed to take $\sim 10^7 \text{ yr}$.

In February 1986, an accurate position was obtained for PSR1855+09 by Fruchter using the Very Large Array (VLA)¹. Shortly afterwards, we began to look for the companion of this pulsar. Observations were made in April 1986 with the 2.3-m telescope at the Wyoming Infrared Observatory. Pictures were taken using the Princeton charged coupled device (CCD) camera at the $f/2$ prime focus. The field of view is $7 \times 10 \text{ arc min}$ and the plate scale is $1.23 \text{ arc s per pixel}$. For the observations of PSR1855+09, 500- and 700-nm bandpass filters were used (each $\sim 100\text{-nm}$ wide), corresponding roughly to the V and R bands.

Nine exposures were made over two nights in the 700-nm band. Of these, the three pictures with the best seeing were averaged and the result used for photometry. Seven 500-nm exposures were taken on one night and the best two averaged. All exposures were 300 s long. Sky transparency was good during the observations. The seeing in the best 700-nm frames is 2.5 arc s (full width at half maximum, FWHM) and 3.0 arc s in the 500-nm frames. The inferior seeing in the 500-nm pictures is due in part to strong, gusty winds during those exposures. The telescope was moved slightly between exposures to avoid always putting

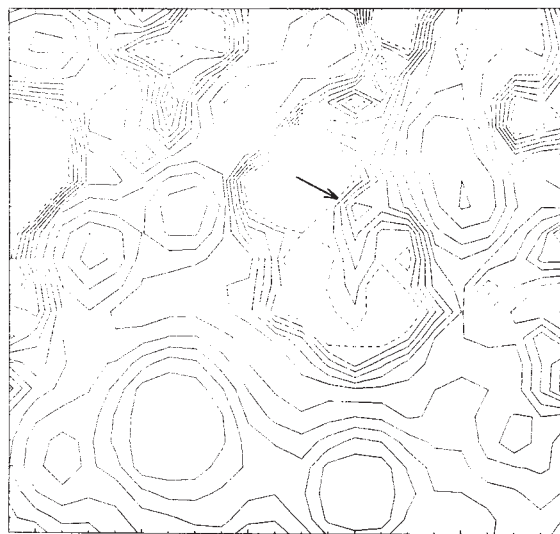


Fig. 1 Isophot plot of the 700-nm-CCD frame of the companion of PSR1855+09. East is up and north is to the right. The arrow indicates the position of the object. The area shown is $27 \times 27 \text{ arc s}$. The contours are linearly spaced for low intensities changing to logarithmic at high intensities. The lowest intensity contour is dashed. The west pointing 'tail' near the marked object is caused by a very faint ($m_{700} = 24$) star visible in the best seeing 700-nm frame. The two bright objects at the western edge of the field are $m_{700} = 17.5$ stars visible on Palomar Sky Survey plates.

the candidate star close to any defective pixels on the CCD.

The candidate star for the companion of PSR1855+09 is shown in Fig. 1. Its peak intensity in the 700-nm frame shown is 7σ above the noise (detector noise plus sky noise). In the poorer seeing 500-nm frame, the peak is 3σ above the noise. The indicated star corresponds, within astrometric error, to the VLA position of PSR1855+09 listed in Table 1. (Error in astrometry is estimated to be $\pm 0.9 \text{ arc s}$ in both right ascension and declination.) It is in a dense field of stars, only 3° away from the galactic plane. The number of stars in the field was

Table 1 Parameters of PSR1855+09

Right ascension (1950.0)*	$\alpha = 18\text{ h } 55\text{ min } 13.681\text{ s} \pm 0.006\text{ s}$
Declination (1950.0)*	$\delta = +09^\circ 39' 13.39'' \pm 0.09''$
Pulsar period*	$P = 5,362,100,452.5 \pm 0.3\text{ ps}$
Orbital period*	$P_b = 1,065,067.49 \pm 0.03\text{ s}$
Dispersion measure distance*	$d = 350\text{ pc}$
Companion mass*	$0.2\text{--}0.4 M_\odot$
Pulsar magnetic field*	$B = (3.4 \pm 0.4) \times 10^8\text{ G}$
Companion apparent magnitude	$m_{700} = 22.01 \pm 0.10$ $m_{500} = 22.65 \pm 0.15$
Companion temperature	$T = 5,900^{+1,400}_{-1,000}\text{ K}$

* Data from ref. 1.

counted to estimate the probability of a chance coincidence. The measured density of stars gives the probability for the centroid of any star to be within 1 pixel of the indicated position of $\leq 5\%$.

Data reduction and photometry were done using the Princeton Stargaze picture processing system. Photometry is complicated by the fact that the candidate star lies in the wings of a pair of nearby bright stars. These stars were fitted by gaussians and subtracted from the pictures. A 3.7×3.7 arc s aperture was centred on the candidate and its intensity measured. A relatively small aperture was chosen to exclude light from other nearby stars. The light falling outside the aperture was estimated by photometry of calibration stars and the intensity was corrected for this. The resulting intensities were converted into AB_{79} system magnitudes using the flux standard stars BD+26 2606 and HD84937 (ref. 3). (AB_{79} magnitudes are relative to a fictitious star whose spectrum is flat in frequency. $m_{AB} = 0$ corresponds to a flux density of $3.65 \times 10^{-20}\text{ erg s}^{-1}\text{ cm}^{-2}\text{ Hz}^{-1}$.) The apparent magnitudes are listed in Table 1. Most of the error is due to the uncertainty in the background level of such a crowded field of stars. Other errors come from sky and readout noise and possible errors in centring the photometry aperture. The colour index $m_{500} - m_{700}$ is 0.6 ± 0.2 .

The candidate star must be a degenerate star. Main-sequence stars with the observed flux ($m_{700} = 22.0$) in the galactic plane are redder than $m_{500} - m_{700} = 2.3$, because intrinsically blue stars must be distant and highly reddened. Degenerate stars at the same apparent magnitude are ~ 100 times closer and therefore suffer less reddening. The probability of a degenerate star being with 1 pixel of the pulsar position is $< 10^{-4}$ (ref. 4). Therefore we conclude that the star indicated in Fig. 1 is the companion of PSR1855+09.

The temperature of the companion star can be estimated from the colour index. Using the dispersion measure distance obtained by Segelstein *et al.*¹ of 350 pc and taking a value for interstellar absorption of $A_V = 1.9\text{ mag kpc}^{-1}$ (ref. 5), we get a corrected colour index $m_{500} - m_{700} = 0.3 \pm 0.2$. Dispersion measure distances are typically accurate to 50%, although in some cases they may be off by a factor of two or more⁶. (The error given includes a 50% uncertainty in distance.) The corrected colour index corresponds to a temperature of $5,900^{+1,400}_{-1,000}\text{ K}$.

From pulse arrival time measurements, Segelstein *et al.*¹ have estimated the mass of the companion of PSR1855+09 to be $0.2\text{--}0.4$ solar masses. The observed object is consistent with a white dwarf of $M = 0.3 M_\odot$, $R = 0.02 R_\odot$, and $T = 5,900\text{ K}$ at a distance of 475 pc, which agrees with the 350 pc dispersion measure distance.

That the companion star is so cool indicates that the PSR1855+09 system is very old. The characteristic cooling time for a $0.3 M_\odot$ white dwarf is $2 \times 10^9\text{ yr}$ (ref. 7). The age of the pulsar, whose progenitor evolves off the main sequence earlier than the white dwarf's, must be greater than this. The pulse period of PSR1855+09 is 5.4 ms. An old pulsar spinning so fast must have been spun up by angular momentum transfer from the companion. Our evidence supports the conventional

hypothesis that millisecond pulsars are recycled, that is they are old neutron stars whose large angular momentum is due to mass accreted from a companion star.

A similar low-temperature white dwarf was observed in the PSR0655+64 system by Kulkarni⁸. The existence of such an old yet still active pulsar is at odds with the naive notion of magnetic field decay in pulsars⁹, which would switch them off after $\tau_{\text{magnetic}} \approx 10^7\text{ yr}$. Kulkarni proposes that the pulsar magnetic field has two components, one that decays in a timescale of 10^7 yr and a residual nondecaying field. Under some circumstances, the residual field may be sufficient for the pulsar to remain active for much longer than τ_{magnetic} .

Measurement of the pulse period derivative yields an estimate of the surface magnetic field of a pulsar¹⁰. The timing data¹ for PSR1855+09 gives a field of $(3.4 \pm 0.4) \times 10^8\text{ G}$. Because the age of the pulsar is 100 times the field decay time, the present field is a good candidate for a nondecaying remnant field. The other millisecond pulsars, PSR1937+21 and PSR1953+29, have surface fields within a factor of two of this value. These three pulsars have the smallest magnetic fields measured. Whether this represents a lower limit on pulsar magnetic fields is an interesting question awaiting more data.

We thank the director and staff of the Wyoming Infrared Observatory for telescope time and assistance. Earl Spillar helped us make the observations. Andy Fruchter, Lloyd Rawley, Dan Stinebring and Joseph Taylor supplied much useful information and made many helpful comments. This work was supported in part by the NSF.

Received 11 June; accepted 29 August 1986.

1. Segelstein, D. J., Rawley, L. A., Stinebring, D. R., Fruchter, A. S. & Taylor, J. H. *Nature* **322**, 714–717 (1986).
2. van den Heuvel, E. P. J. & Tamm, R. E. *Nature* **309**, 235–237 (1984).
3. Oke, J. B. & Gunn, J. E. *Astrophys. J.* **266**, 713–717 (1983).
4. Liebert, J. A. *Rev. Astr. Astrophys.* **18**, 363–398 (1980).
5. Allen, C. W. *Astrophysical Quantities* 3rd edn, 263 (Athlone, London, 1973).
6. Lyne, A. G., Manchester, R. N. & Taylor, J. H. *Mon. Not. R. astr. Soc.* **213**, 613–639 (1985).
7. Shapiro, S. L. & Teukolsky, S. A. *Black Holes, White Dwarfs and Neutron Stars*, 86 (Wiley, New York, 1983).
8. Kulkarni, S. R. *Astrophys. J.* **306**, L85–L89 (1986).
9. Bhattacharya, D. & Srinivasan, G. *Curr. Sci.* **55**, 327–330 (1986).
10. Manchester, R. N. & Taylor, J. H. *Pulsars*, 177 (Freeman, San Francisco, 1977).

An explanation for fast downflows on the Sun

A. N. McClymont* & I. J. D. Craig†

* Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA

† Department of Mathematics, University of Waikato, Private Bag, Hamilton, New Zealand

The discovery of coronal loops has provoked major revisions in the theory of solar and stellar atmospheres. The solar corona is now recognized as a highly structured plasma, organized by bundles of intense magnetic field into individual arch- and loop-like structures along which gas can flow¹. Each arch, rooted in the solar interior, extends through the chromosphere ($T < 10^4\text{ K}$), and generally contains coronal plasma ($T > 10^6\text{ K}$) in its upper reaches. Early attempts at explaining the temperature and density structure of isolated static loops met with considerable success, despite the simplicity of the theory^{2,3}; however, attempts to generalize the theory to describe loops with mass flows have remained singularly unsuccessful^{4–9}. Here we point out that the puzzling observational properties of loop mass flows^{10–15} can be naturally explained if the loop is sufficiently cool ($T \lesssim 10^6\text{ K}$)—that is, if it contains no coronal plasma.

Observed flows exhibit several features that must be explained by any convincing theory. First, the flows are observed to persist much longer than the transit time of a gas element through the

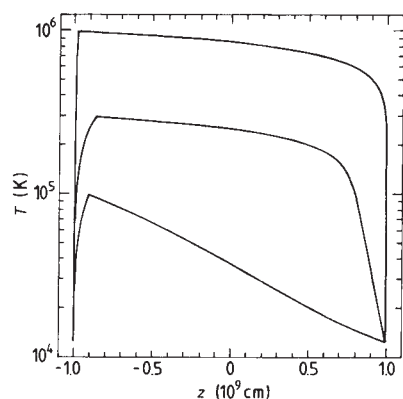


Fig. 1 Temperature profile in a loop of length 2×10^9 cm (z is the distance along the loop) with a steady flow of 5 km s^{-1} at 10^5 K, for peak temperatures of 10^5 , 3×10^5 and 10^6 K. The corona is heated by a point energy source (at the position of maximum temperature). Extreme asymmetry is required to drive flows of the observed magnitude, $5\text{--}10 \text{ km s}^{-1}$ at 10^5 K.

loop^{10,11}. This fact alone suggests a quasi-steady model in which plasma flows continually from one footpoint to the other. This model also readily accounts for the second observed property, namely, that emission from the downflowing leg predominates¹⁰⁻¹⁵. This can be understood as a consequence of the asymmetry of the temperature profile⁹ induced by steady flow (see Fig. 1). Altogether more tantalizing, however, is the third property: flow speeds are large ($5\text{--}10 \text{ km s}^{-1}$) at transition-region temperatures ($T \approx 10^5$ K); yet the flows apparently disappear at coronal temperatures¹¹⁻¹⁴. This feature appears completely inexplicable in terms of current steady flow models. The problem arises because, on the assumption of subsonic flow, the pressure $P = 2\rho RT$ should be nearly uniform, which, together with the steady-state condition of constant mass flux $q = \rho v$, implies $v \propto T$. It follows that in typical loops the highest velocities are to be expected at coronal temperatures, namely $T \approx 10^6$ K.

The key to this paradox lies in momentum conservation. When the flow velocity approaches the sound speed, the dynamic pressure, $\rho v^2 \equiv qv$, becomes important, forcing the gas pressure to decrease. The net result is an upper limit on the coronal temperature for a given flow speed (see equation (3) below).

To illustrate the problem in the simplest possible way, we neglect gravity and assume a loop of uniform area. Then the steady flow momentum equation implies

$$qv + P = qv_0 + P_0 \quad (1)$$

where subscript zero denotes conditions at a reference point. As the dynamic pressure qv increases, the gas pressure P must decrease. Eliminating $\rho = q/v$ from equation (1) and writing $M = v/c$ for the Mach number relative to the isothermal sound speed $c = \sqrt{2RT}$, we can solve for temperature, obtaining

$$\frac{T}{T_0} = \frac{v}{v_0} \left[(1 + M_0^2) - M_0^2 \frac{v}{v_0} \right] \quad (2)$$

For subsonic flow ($v \ll c$), equation (2) reduces to the familiar hydrostatic condition of uniform pressure, $v/v_0 = T/T_0$. But at higher flow speeds the function $T(v)$ is a parabola, with a turning point at (v^*, T^*) , where

$$\frac{v^*}{v_0} = \left(\frac{1 + M_0^2}{2M_0^2} \right); \quad \frac{T^*}{T_0} = \left(\frac{1 + M_0^2}{2M_0} \right)^2 \quad (3)$$

which occurs when the flow speed reaches the isothermal sound speed.

This turning point defines the maximum possible temperature T^* in the loop. Setting $T_0 = 10^5$ K and noting that for $v_0 \leq 10 \text{ km s}^{-1}$, $M_0^2 \ll 1$, we obtain from equation (3):

$$T^* \approx (20/v_0)^2 \times 10^5 \text{ K} \quad (4)$$

with v_0 in km s^{-1} . It is clear that higher loop temperatures correspond to lower flow speeds. For v_0 in the range $5\text{--}10 \text{ km s}^{-1}$, equation (4) yields $T \approx 4 \times 10^5\text{--}1.6 \times 10^6$ K. Thus the absence of significant flow velocities at coronal temperatures is a natural consequence of momentum conservation.

So far the cause of mass flow has not been considered. A complete description of the flow requires consideration of the energy equation^{9,16-18}; in the present model, spatial asymmetry of the coronal energy source powers the flow. Figure 1 shows temperature profiles for peak temperatures of 10^5 , 3×10^5 and 10^6 K, with a flow speed of 5 km s^{-1} at 10^5 K, computed for the simplest case of a localized 'point' energy source. The heating asymmetry is immediately apparent. Indeed, extreme heating asymmetry is required to drive flows of the observed magnitude. The flow originates as follows⁹. Energy is transported most readily (by thermal conduction) towards the closer footpoint, where the excess heat flux 'boils off' chromospheric material. The conductive flux into the downflowing leg of the loop is weaker, allowing the gas to cool as it flows towards the distant footpoint, where it condenses back into the chromosphere. Note that the coronal heat source is only $\sim 10\%$ of the loop length above the chromosphere in the upflowing leg. Larger flow speeds require even greater asymmetries¹⁸.

Turning briefly to the question of gravity and area non-uniformity, we find that these can significantly alter the flow pattern, but in opposite senses. Gravity causes the gas density to decrease with increasing height; thus the velocity of upflowing plasma, $v = q/\rho$, tends to increase more quickly, enhancing the influence of dynamic pressure and reducing the maximum temperature (equation (3)) still further. Fast flows may therefore be constrained not only to cool loops, but also to low-lying loops. If the loop area A increases with height, the velocity, $v = q/A\rho$, rises less quickly, so that the constraints on loop temperature and height are somewhat relaxed. However, physically realistic area variations seem unlikely to produce a major effect¹⁸.

Thus we conclude that all features of the observed flows follow from the analysis presented here. Specifically, steady flows that exhibit substantial transition region velocities with a predominance of redshifted emission, and which disappear at coronal temperatures, are naturally explained. The theory will be strengthened if spatially resolved observations can confirm the predicted asymmetry in loop structure associated with fast flows.

This research was partially supported by NSF grant ATM85-96006.

Received 24 March; accepted 19 June 1986.

- Withbroe, G. L. in *Solar Active Regions* (ed Orrall, F. Q.) 199-212 (Colorado Associated University Press, Boulder, 1980).
- Craig, I. J. D., McClymont, A. N. & Underwood, J. H. *Astr. Astrophys.* **70**, 1-11 (1978).
- Rosner, R., Tucker, W. H. & Vaiana, G. S. *Astrophys. J.* **220**, 643-655 (1978).
- Yeh, T. *Sol. Phys.* **55**, 241-250 (1977).
- Pneuman, G. W. & Kopp, R. A. *Sol. Phys.* **57**, 49-64 (1978).
- Cargill, P. J. & Priest, E. R. *Sol. Phys.* **65**, 251-269 (1980).
- Athay, R. G. *Astrophys. J.* **249**, 340-348 (1981).
- Noci, G. *Sol. Phys.* **69**, 63-76 (1981).

- Craig, I. J. D. & McClymont, A. N. *Astrophys. J.* **307**, 367-380 (1986).
- Gebbie, K. B. *et al. Astrophys. J.* **251**, L115-L118 (1981).
- Feldman, U., Cohen, L. & Doschek, G. A. *Astrophys. J.* **255**, 325-328 (1982).
- Dere, K. P. *Sol. Phys.* **77**, 77-93 (1982).
- Doschek, G. A., Feldman, U. & Bohlin, J. D. *Astrophys. J.* **205**, L177-L180 (1976).
- Lites, B. W. *et al. Astrophys. J.* **210**, L111-L113 (1976).
- Athay, R. G., Gurman, J. B., Henze, W. & Shine, R. A. *Astrophys. J.* **265**, 519-529 (1983).
- Boris, J. P. & Mariska, J. T. *Astrophys. J.* **258**, L49-L52 (1982).
- Mariska, J. T. & Boris, J. P. *Astrophys. J.* **267**, 409-420 (1983).
- McClymont, A. N. & Craig, I. J. D. *Astrophys. J.* (in the press).

Poly(styrene peroxide): an auto-combustible polymer fuel

K. Kishore & T. Mukundan

Department of Inorganic and Physical Chemistry, Indian Institute of Science, Bangalore 560 012, India

Commonly, polymers exhibit extremely slow, candle-like burning, although oxidizers can increase the rate of combustion, as in solid propellants. Polymer degradation occurs before combustion, and being endothermic, demands a continuous supply of energy. However, if a polymer can degrade exothermically, it can sustain degradation and automatically increase the rate of combustion (autopyrolysis). Poly(styrene peroxide), PSP, which degrades exothermically¹ is such an exceptional and rare example. We explored the possibility of autopyrolytic behaviour in PSP. Interestingly, it was observed that, once initiated, PSP not only sustains its own degradation, but also supports a flame in air, with a burning rate comparable to that of solid propellants. This makes PSP an ideal candidate for a special fuel, such a fast-burning autopyrolytic polymer being hitherto unknown. Here, in the first reported study on the combustion of a polymeric peroxide, we examine some salient features of the combustion of this unique auto-combustible polymer and potential fuel, and show that the combustion is controlled by polymer degradation.

PSP was prepared and purified from styrene and oxygen as reported by Miller and Mayo². It was characterized by ¹H and ¹³C NMR, IR and Raman spectroscopy. The PSP samples containing additives (nitrobenzene and *p*-toluidine) were prepared by dissolving PSP and the additive (one mole for every 20 polymer repeat units) in acetone, followed by removal of solvent. The polymer was filled free of air bubbles in glass tubes 8 mm in diameter with cotton support. The sample was ignited using a nichrome wire, and the linear burning rate was measured. The burning rates (the average of five repetitions) were closely reproducible (within 1% error). Temperature profile measurements were carried out using a chromel-alumel thermocouple (0.127 mm in diameter) inserted into holes drilled into the polymer sample so that the thermocouple bead was well inside. The thermocouple output was fed to a Philips (PM 3234, 0–10 MHz) double beam cathode ray storage oscilloscope, maintained at a sensitivity of 5.8 mV cm⁻¹. The sample tubes were insulated with plaster of paris to minimize lateral heat losses. The degradation peak temperatures of PSP with and without additives were obtained on a DuPont 990 differential scanning calorimeter at a heating rate of 20 °C per min with a sample weight of ~2 mg. The temperature axis was calibrated using benzoic acid.

Pyrolysis gas chromatography studies of the same samples were carried out on a Shimadzu GC-4C gas chromatograph, with attached PYR-1A pyrolyser using a dioctyl phthalate coated on a chromosorb column. Nitrogen (35 ml min⁻¹) was used as carrier gas. The samples were pyrolysed at 200 °C, and chromatographed on the column heated at 5 °C min⁻¹, from 50–125 °C. The products were identified by comparison with retention times of pure compounds. The three sample weights varied and hence the aldehyde peak intensities are not the same.

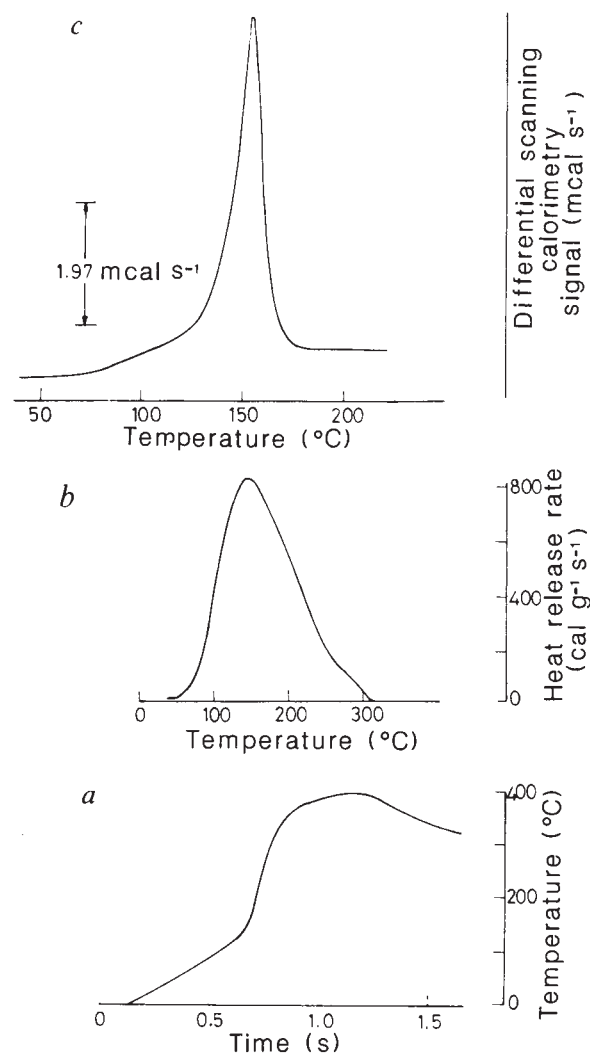


Fig. 1 Thermal profile of poly(styrene peroxide). *a*, Time-temperature trace during combustion, *b*, computer generated, rate of heat release against time plot, *c*, differential scanning calorimetry trace of degradation; heating rate = 20 °C min⁻¹, sample weight ~2 mg.

PSP degrades via a radical chain scission mechanism³ giving benzaldehyde and formaldehyde as major products. A recent pyrolysis gas chromatography study (in the pyrolysis range 120–430 °C) confirms the formation of these two products and accounts for more than 90% of the total degradation products⁴. The remaining products might be formed due to the radical and ionic disproportionation mechanism as proposed by Mayo and Miller³.

A typical temperature profile (time-temperature trace) is shown in Fig. 1*a*. Heat release rate (dq/dt) at any instant can be calculated from the equation,

$$\frac{dq}{dt} = C_p \frac{T_2 - T_1}{t_2 - t_1}$$

T_2 and T_1 being the temperatures at time t_2 and t_1 respectively, and C_p the average heat capacity of PSP. The profile data were fed into the computer to obtain a plot of dq/dt against T . The temperature dependence of C_p , measured by differential scanning calorimetry was obtained using a computer fit ($C_p = 0.001236T + 0.3541$). The value of C_p above 90 °C, when PSP starts to degrade during combustion, was determined by extrapolation of this fit. The plot of dq/dt against T is given in Fig. 1*b*. A typical calorimetric trace (of PSP degradation) that relates the rate of heat flow (dq/dt) to time/temperature, is shown in

Table 1 Degradation peak temperatures and burning rates of PSP with and without additives

Additives (5 mole % per repeat unit of PSP)	DSC peak temperature (°C)	Linear burning rate (mm s ⁻¹)
Nil	149.0 ± 0.2	2.00 ± 0.03
<i>p</i> -Toluidine	141.8 ± 0.2	2.47 ± 0.03
Nitrobenzene	158.4 ± 0.2	0.162 ± 0.001

DSC, differential scanning calorimetry.

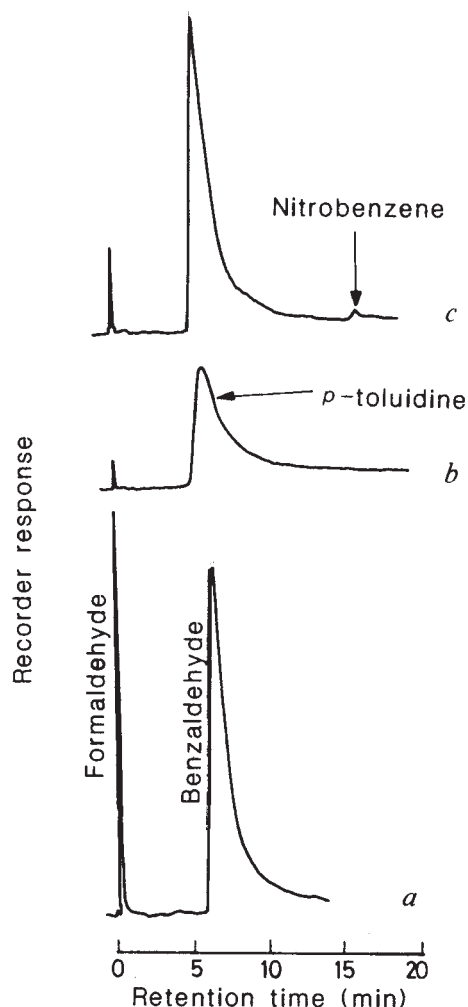


Fig. 2 Pyrograms of PSP samples. *a*, PSP, *b*, PSP + *p*-toluidine (merges with benzaldehyde) and *c*, PSP + nitrobenzene.

Fig. 1c. Note that the solid-gas interface temperatures (surface temperatures) where PSP degradation occurs coincides in the two cases shown in Fig. 2b, c.

The profile experiment was used to determine the heat of degradation of PSP during its combustion. PSP initially at ambient temperature degrades and the gaseous products are heated to a maximum temperature before entering the flame at the mouth of the tube. The enthalpy change for PSP going from ambient to maximum gas temperature is $\Delta H = C_p(T_2 - T_1)$, T_2 and T_1 being the maximum gas temperature (400 °C) in the profile (Fig. 1a) and ambient temperature (25 °C) respectively; an average value of $C_p = 0.54 \text{ cal mol}^{-1}$ at the surface temperature, 150 °C, was used. Corrections for vapourization of benzaldehyde and formaldehyde¹ were made to the above value of ΔH and the heat of degradation of PSP from the combustion experiment thus obtained was $45.0 \text{ kcal mol}^{-1}$. Considering the heat losses that occur during the combustion experiment this value is in reasonable agreement with the experimental value reported earlier¹. Comparing the heat of degradation value obtained from the profile experiment and those from differential scanning calorimetry ($50 \pm 1 \text{ kcal mol}^{-1}$) and thermochemical calculations¹ ($52 \pm 1.5 \text{ kcal mol}^{-1}$) suggests that combustion of PSP occurs as a consequence of its degradation, otherwise heat released before combustion should differ from the heat of degradation obtained by scanning calorimetry. This conclusion is supported by the following observations. (1) During combustion, the flame remains at the mouth of the tube, while the surface at which degradation occurs recedes, that is, the flame persists as a result of the combustion of PSP degradation prod-

ucts; (2) under controlled conditions PSP degrades without a flame, the regression rate being equal to the burning rate.

A corroboration of this inference was sought with the help of two additives. Thermograms were recorded to obtain degradation peak temperatures with and without these additives; the linear burning rates were also measured (Table 1). It can be seen that the additives bring about similar changes in combustion and degradation; *p*-toluidine accelerates both, while nitrobenzene decelerates both.

Pyrolysis gas chromatography of the PSP samples containing additives was carried out to investigate the effect on PSP degradation (Fig. 2). No new degradation products were formed in the presence of the additives and the overall composition of the degradation products remained the same. In the case of the PSP-nitrobenzene sample, nitrobenzene has a longer retention time (Fig. 2c), while for the PSP-*p*-toluidine case the *p*-toluidine merges with the benzaldehyde peak (Fig. 2b). These results suggest that the additives only affect the rate of reaction rather than the mechanism of PSP degradation. A detailed study of this aspect, which unavoidably falls beyond the realm of this preliminary report, would prove vital in regulating the rate of combustion for potential use of the polymer as a fuel. However, the present investigation describes the unusual phenomena of autopyrolysis and autocombustion of an unadulterated polymer and unambiguously proves that its combustion is primarily controlled by degradation.

We thank Mr P. A. Joy and Mr M. Jayamurthy for their help in computations.

Received 17 March; accepted 5 September 1986

1. Kishore, K. & Ravindran, K. *Macromolecules* **15**, 1638-1639 (1982).
2. Miller, A. A. & Mayo, F. R. *J. Am. Chem. Soc.* **78**, 1017-1023 (1956).
3. Mayo, F. R. & Miller, A. A. *J. Am. Chem. Soc.* **78**, 1023-1034 (1956).
4. Kishore, K. & Ravindran, K. *J. analyt. appl. Pyrol.* **5**, 363-370 (1983).

Hydrothermal methane plumes in the Mariana back-arc spreading centre

Y. Horibe*‡, K.-R. Kim†‡ & H. Craig†

* Ocean Research Institute, University of Tokyo, Nakano, Tokyo 164, Japan

† Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093, USA

Submarine hydrothermal vents were first unequivocally detected by measurement of $^3\text{He}/^4\text{He}$ ratio anomalies coupled with deviations from ambient T - S relationships on the Galapagos spreading centre^{1,2}. Since then the existence of hydrothermal vents along the entire length of the East Pacific Rise has been demonstrated by shipboard observations of ^3He , methane, and manganese in the effluent plumes³⁻⁵. We have extended our studies of vent fields at mid-ocean-ridge spreading centres to search for effluent plumes in back-arc basins, which are also thought to be active regions of seafloor spreading where new oceanic crust is formed⁶, and thus where rifting and fissuring along the axis of extension should provide similar opportunities for the downward penetration of sea water and the development of active hydrothermal systems. We have detected large plumes of methane-enriched water in the Mariana Trough back-arc basin⁷ and also in the summit crater of Loihi Seamount⁷ at the present site of the Hawaiian hotspot, thus establishing the widespread existence of 'off-ridge' hydrothermal vents in the ocean. However, although the Loihi vents produce plumes enriched in both ^3He and methane⁷, we find that, surprisingly, the Mariana vents are enriched in methane without a corresponding enrichment in ^3He . This is the first time such a 'decoupling' of these two effluent plume tracers has been observed.

‡ Present addresses: Scripps Institution of Oceanography, La Jolla, California 92093, USA (Y.H.); Department of Oceanography, Seoul National University, Seoul 151, Korea (K.-R.K.).

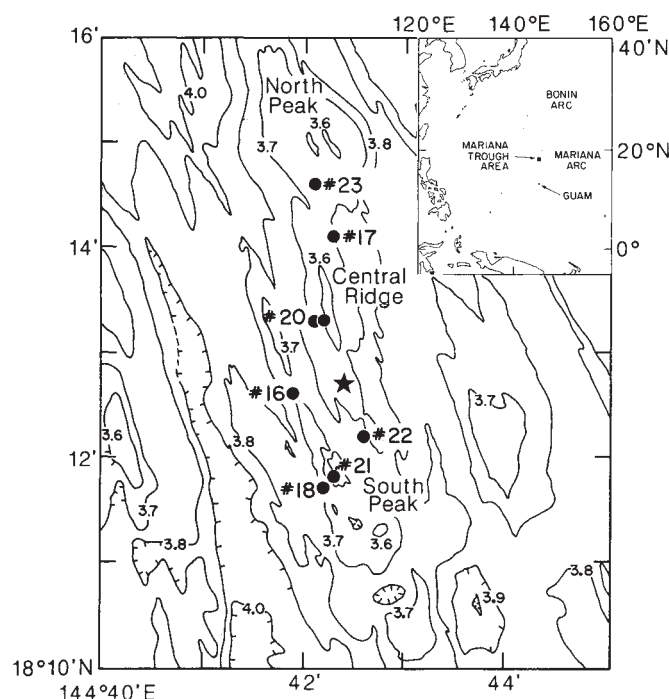


Fig. 1 Cepheus station locations in the Mariana Trough. The stations lie between two large volcanic peaks four miles apart, along a ridge of fresh lavas from fissure eruptions. This section is a high spot in the Mariana Trough and was chosen for investigation because the axis is accessible for study by the submersible Alvin (depth < 4 km). The star on Central Ridge marks the site of a 0.02 °C temperature anomaly observed with Deep Tow by P. Lonsdale, who made this map. A CTD towed on the Cepheus expedition also detected a 'true' hydrothermal plume spike in the vicinity of station 20, with amplitude 0.01 °C. Depth contours in km.

Our search for off-ridge hydrothermal vents was carried out in a region of the axial rift valley between 18°11' N to 18°15' N, at the suggestion of P. Lonsdale, who had previously (with the Deep Tow CTD (conductivity, temperature, depth) observed small temperature spikes in the water column that were 'real' plumes in the sense that corresponding decreases in salinity were lacking¹. The shipboard work was done on the Japanese vessel *Hakuho Maru* on the Cepheus expedition⁸, in the area shown in Fig. 1, which was mapped by Lonsdale. Seven hydrographic stations were occupied along the axial volcanic ridge; a methane plume was observed in all cases although complete profiles through the plume were obtained at only four of the stations.

Figure 2 shows the Mariana methane plume in section from south to north along the ridge crest. The core concentration of methane is $74 \times 10^{-6} \text{ cm}^3 \text{ kg}^{-1}$ at a depth of 2950–3000 m, increasing slightly toward the northern peak. The plume rises 600–700 m above the bottom before entraining enough ambient water to reach density equilibrium and spread horizontally along the 3000 m isopycnal ($1041.428 \text{ kg m}^{-3}$ at 2957 m). This 'upper' plume is distinctly separated from the lower more discontinuous plume ~200 m above the sea floor, and it is obviously generated by more active vents. We suppose that the 'upper' plume is a '21° N'-type plume generated by the so-called 'black smokers', which vent high-temperature water from chimneys above the sea floor, whereas the lower and much more dilute plumes probably represent the low-temperature 'Galapagos-type' effluent from a cooled subsurface reservoir. These lower plumes are discontinuous in section because of the offset of the axial ridge at 18°12.5' N (Fig. 1). We have observed⁹ similar sets of near-bottom and 'upper' plumes on the East Pacific Rise at 20° S, where the plumes are clearly associated with vents as shown by Angus photographs of the vent fauna. However, the

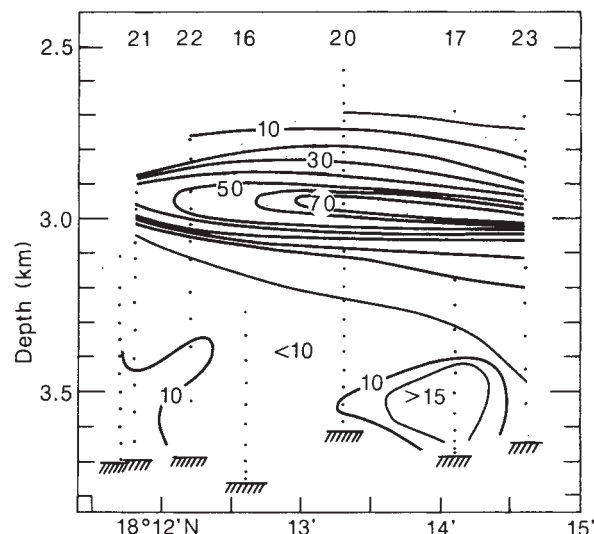


Fig. 2 Structure of the methane plumes along the Mariana Trough spreading axis. Station numbers are along the top; methane contours are in $10^{-6} \text{ cm}^3 \text{ (STP) methane per kg sea water}$. Bottom depths at the stations are indicated by cross-hatched bars.

East Pacific Rise plumes generally rise only some 300 m above the ridge crest before spreading horizontally^{9,10}, much less than the height observed here. As we observed that the potential density gradients are quite similar both in the Mariana plume and at several localities on the East Pacific Rise, it is possible that the much greater rise height of the Mariana plume is due to higher effluent temperatures, perhaps above 400 °C, from very large 'black smokers'.

Table 1 shows the methane and helium data at station 23, the only station for which a helium profile was obtained. The $^3\text{He}/^4\text{He}$ ratio anomalies are tabulated as $\delta^3\text{He}$ values (in per cent), that is, the percentage enrichments relative to the atmospheric ratio. The saturation anomalies relative to atmospheric solubilities¹¹ are also listed for both He isotopes and for neon, and finally the absolute He isotope concentrations are tabulated for each isotope.

Figure 3 shows the $\delta^3\text{He}$ profile superimposed on the methane plumes from stations 20 and 23, and it is seen that the very large and sharp methane maximum is not reflected in the ^3He profile. A close reading of Table 1 shows that there is probably a very slight maximum in $\delta^3\text{He}$ between 3,000 and 3,145 m, correlated with a similar slight maximum in the solubility anomalies and absolute concentrations of the He isotopes. The solubility anomalies and absolute concentrations of ^3He indicate that there may be a small maximum at 3,000 m, amounting to no more than 0.8%. Comparison with the methane peak indicates that the methane/ ^3He ratio in the added volatiles must be $>100 \times 10^{-6}$, which is much greater than the $\sim 3 \times 10^{-6}$ observed at East Pacific Rise vent fields^{5,12}. We also find that, in detail, the ^3He and ^4He concentrations are well correlated ($r=0.87$), with a slope corresponding to a mixture of a He component with a $^3\text{He}/^4\text{He}$ ratio of only 3.2 (± 0.7) times atmospheric, rather than a ratio of 8 as observed on mid-ocean ridges. This indicates that there is a significant radiogenic He component in the Mariana vents in addition to the large methane enrichment relative to He. These data as a whole thus suggest that some effects we have not encountered on the East Pacific Rise have significantly influenced the volatile compositions in the Mariana vents.

A methane profile was also measured in the so-called 'Mounds' area west of the axis of the Mariana Trough at 18°01' N, 144°18' E, an area of high heat flow thought to be a site of hydrothermal activity¹³. Samples were collected at 12 depths from 2,930 to 3,515 m, above the bottom depth of 3,534 m. The observed methane concentrations had a mean of $(4.3 \pm 0.4) \times 10^{-6} \text{ cm}^3 \text{ kg}^{-1}$, which is identical to the ambient methane back-

Table 1 Methane, ^3He , ^4He , Ne and potential temperature (T_{pot}) measurements for Cepheus expedition station 23

Depth (m)	T_{pot} ($^{\circ}\text{C}$)	Methane ($10^{-6} \text{ cm}^3 \text{ kg}^{-1}$)	$\delta^3\text{He}$ (%)	ΔNe (%)	$\Delta^4\text{He}$ (%)	$\Delta^3\text{He}$ (%)	^4He ($10^{-6} \text{ cm}^3 \text{ kg}^{-1}$)	^3He ($10^{-12} \text{ cm}^3 \text{ kg}^{-1}$)
2,705	1.520	4.6	—	—	—	—	—	—
2,803	1.501	8.5	25.9	(3.60)	7.54	37.1	42.46	74.9
2,852	1.486	14.7	27.2	(4.12)	7.99	39.0	42.64	75.9
2,901	1.472	19.4	26.7	(5.18)	7.70	38.1	42.53	75.4
2,949	1.487	45.0	27.4	3.30	7.89	39.1	42.60	76.0
2,998	1.482	71.7	27.6	3.18	8.06	39.5	42.67	76.2
3,047	1.467	33.0	27.7	3.19	8.11	39.7	42.69	76.3
3,145	1.457	16.6	27.8	(3.52)	7.73	39.4	42.54	76.1
3,242	1.457	8.5	27.3	(3.83)	8.12	39.3	42.70	76.1
3,340	1.456	7.9	27.3	3.09	7.80	38.9	42.57	75.9
3,438	1.436	3.4	26.7	3.07	7.86	38.3	42.60	75.6
3,536	1.435	3.5	27.3	3.19	7.99	39.1	42.65	76.0

Seawater samples were collected in 23-litre Niskin hydrographic bottles, with the use of a pinger to measure distance above the sea floor. Samples of 40 ml (He) and 1 liter (methane) were drawn in tandem in copper tubing and Pyrex cylinders. The He samples were sealed with refrigeration clamps and were extracted at Scripps Institution of Oceanography and analysed for $^3\text{He}/^4\text{He}$ ratios and He and Ne concentrations (by isotope dilution) as described by Lupton and Craig¹⁰. Methane measurements were made on board with the Shimadzu FID (flame ionization detection) gas chromatograph and vacuum extraction system described by Kim⁹.

The salinity at 2,705 m is 34.657‰; below this depth the salinity is $(34.663 \pm 0.002)\%$, that is, constant to within the precision of the measurements⁸. $\delta^3\text{He} = (R/R_A - 1) \times 100$, where $R = ^3\text{He}/^4\text{He}$ in the sample and R_A refers to atmospheric $^3\text{He}/^4\text{He}$ ratio ($= 1.40 \times 10^{-6}$). The Δ values for Ne, ^3He and ^4He are the saturation anomalies relative to atmospheric solubility at the potential temperature of the sample¹¹. Slight corrections for air leakage into the copper samplers have been made to the He results for five of the samples in which excess Ne was observed (ΔNe values in parentheses). ΔNe values between 3.0 and 3.3% are normally distributed owing to random errors, and the mean of these values, 3.17%, is taken as the actual Ne solubility anomaly. For the five samples with $\Delta\text{Ne} > 3.3\%$, the ΔHe values are reduced by subtracting 1.276 (the solubility coefficient ratio) times the Ne solubility deviation from the mean. The recalculated anomalies in ^3He and ^4He ratios and solubilities, and the absolute concentrations of ^3He and ^4He , are tabulated for these five samples. The precision of the ^4He concentration data is $\sim 0.3\%$. The precision of the $\delta^3\text{He}$ measurements is about 0.7%, so that the corrections for air leakage for these data are not significant except for a 0.6% correction for the 2,901 m sample. However, these corrections were made very carefully because the lack of a sharp maximum in the isotope ratios or absolute concentrations made us look very closely at the data.

ground in the Mariana Trough samples, $(3.9 \pm 0.6) \times 10^{-6} \text{ cm}^3 \text{ kg}^{-1}$, obtained by averaging 23 samples with minimum methane concentrations in the same depth range at stations 16–23 on the axis of the trough. Thus if hydrothermal activity is present at the 'Mounds' site, it is not intense enough to make observable methane plumes and it must be extremely diffuse.

In conclusion, we have observed a large hydrothermal methane plume rising 700–800 m above the sea floor and extending several miles along the axial region of the Mariana Trough. The discovery of this plume extends the source of submarine hydrothermal vents to the extensional region of back-arc basins, which implies that there is much more hydrothermal activity in

the ocean than previously thought. However, the lack of a strong helium signal associated with the methane plume, and the relatively low $^3\text{He}/^4\text{He}$ ratio in the added He, indicate that significant differences in chemistry are present in the mantle sources and hydrothermal water-rock interaction processes in the marginal basins of the oceans. Thus we look to the forthcoming Alvin investigations in these regions to increase our understanding of these effects.

We thank H. Tsubota, M. Fukazawa, T. Gamo, H. Inoguchi, H. Morinaga and the officers and crew of the vessel for assistance aboard the *Hakuho Maru*. We are grateful to Peter Lonsdale for his detailed map and discussions of the Mariana Trough area, and to R. L. Chase for helpful comments. The work at La Jolla was supported by the NSF Division of Ocean Sciences. *Note added in proof:* Since this paper was submitted we have extended our observations of back-arc hydrothermal methane plumes to the central spreading axis of the Lau Basin in the southwest Pacific¹⁴. Hawkins and Helu¹⁵ have dredged sulphides of probable hydrothermal origin from this basin, and Both *et al.*¹⁶ have photographed hydrothermal chimneys and vent fauna on the spreading axis of the Manus Basin, where we subsequently observed large methane plumes (Papatua expedition, 1986).

Received 6 May; accepted 13 August 1986.

- Weiss, R. F., Lonsdale, P., Lupton, J. E., Bainbridge, A. E. & Craig, H. *Nature* **267**, 600–603 (1977).
- Lupton, J. E., Weiss, R. F. & Craig, H. *Nature* **267**, 603–604 (1977).
- Lupton, J. E. *et al. Earth planet. Sci. Lett.* **50**, 115–127 (1980).
- Craig, H. *Eos* **62**, 893 (1981).
- Kim, K.-R., Craig, H. & Horibe, Y. *Eos* **64**, 724 (1983).
- Karig, D. E. *J. geophys. Res.* **76**, 2542–2561 (1971).
- Horibe, Y., Kim, K.-R. & Craig, H. *Eos* **64**, 724 (1983).
- Horibe, Y. *Prelim. Rep. CEPHEUS Exped. 1982*, 1–78 (Ocean Research Inst., Tokyo, 1983).
- Kim, K.-R. Ph.D. thesis, Univ. California at San Diego (1983).
- Lupton, J. E. & Craig, H. *Science* **214**, 13–18 (1981).
- Weiss, R. F. *J. chem. Engng Data* **16**, 235–241 (1971).
- Welhan, J. A. & Craig, H. *Geophys. Res. Lett.* **6**, 829–831 (1979).
- Lonsdale, P. & Hawkins, J. *Bull. geol. Soc. Am.* **96**, 940–951 (1985).
- Craig, H., Kim, K.-R. & Poreda, R. *Eos* **67**, 378 (1986).
- Hawkins, J. & Helu, S. *Eos* **67**, 378 (1986).
- Both, R. *et al. Eos* **67**, 489–490 (1986).

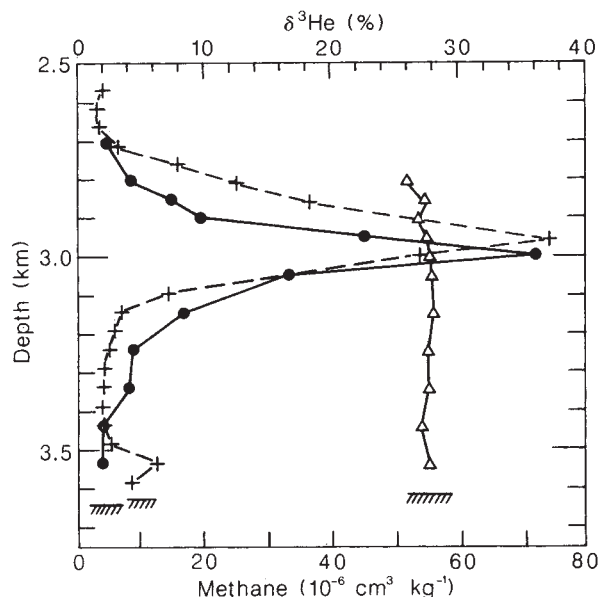


Fig. 3 Methane plumes at stations 20 (dashes) and 23 (solid), compared with the $\delta^3\text{He}$ isotopic ratio anomaly relative to atmospheric helium at station 23 (triangles).

Structure of the Kerguelen Plateau province from Seasat altimetry and seismic reflection data

Millard F. Coffin*, Hugh L. Davies*
& William F. Haxby†

* Bureau of Mineral Resources, Geology and Geophysics,
GPO Box 378, Canberra, A.C.T. 2601, Australia

† Lamont-Doherty Geological Observatory, Palisades,
New York 10964, USA

The Kerguelen Plateau in the southern Indian Ocean stands 2–4 km above the adjacent sea floor and is ~2,500 km long (Fig. 1)^{1,2}. Seasat provided a unique data set for deriving the free-air gravity field of the region (Fig. 2)^{3–5}. Here we report the results of a new analysis of the plateau province's structure employing both Seasat and newly-acquired multichannel seismic (MCS) data^{6,7} for ground truth. The northern sector is characterized by volcanism and a sedimentary basin; the southern sector by a broad anticlinal arch, major faulting, and a sedimentary basin; and the eastern sector by an abyssal basin (Labuan) and bounding ridge (William's). The three sectors argue for a more complex tectonic evolution of the feature than has been previously proposed.

Seasat's radar altimeter, pulsing 1,000 times per second, allowed the determination of the average distance between the satellite and several square kilometres of the sea surface directly beneath it. One-second averages of the distances, each representing ~8 km of ground track, provided continuous profiles of sea surface topography to a precision of 5–10 cm. The ocean surface closely coincides with the geoid, the gravitational equipotential surface reflecting the mass distribution within the Earth. Variations in gravitational acceleration computed directly from the geoid surface⁴ result in gravity anomaly estimates accurate to

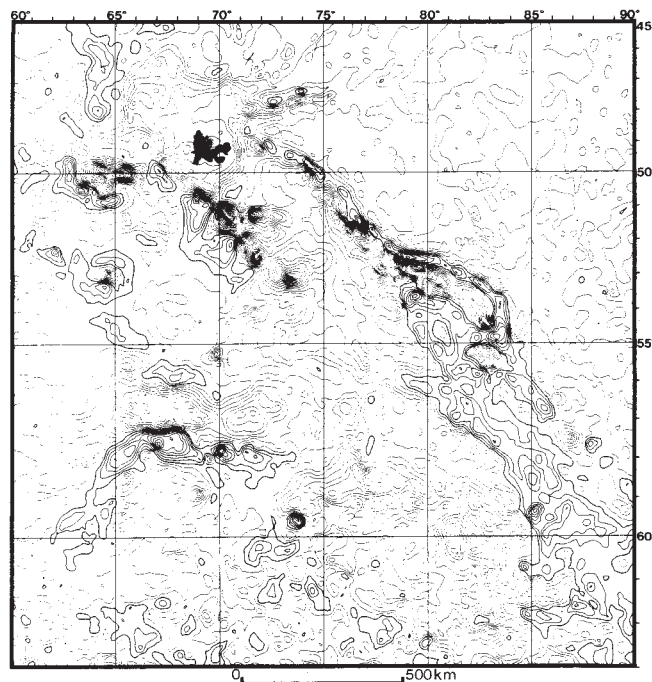


Fig. 2 Seasat-derived free-air gravity field of the Kerguelen Plateau. The contour interval is 8 mGal, and the 0 mGal contour is drawn in a heavy black line. Kerguelen and Heard islands are in solid black. Compare with bathymetry in Fig. 1.

5 mGal with a spatial resolution of 25–50 km. Although less accurate than shipboard gravity measurements at spatial scales <100 km, Seasat's uniform track coverage (<5 to several tens of kilometres spacing)⁵ permits comprehensive mapping of the gravity field with consistent accuracy (Fig. 2), whereas shipboard measurements are sparse in the plateau region (Fig. 1) and are inadequate for regional analyses. Caution must be exercised in viewing the gravity map: south of ~60°S sea ice increased altimeter noise by an order of magnitude and thus reduced precision; features of dimensions less than several tens of kilometres may be artefacts of the sea-surface gridding procedure or represent measurement noise; and features lineated in the satellite track directions may be artificial, resulting from uncertainties in correlating measurements along adjacent profiles.

In 1985, the R/V *Rig Seismic* acquired the first multichannel seismic data over the southern and eastern sectors of the Kerguelen Plateau province^{6,7}. Two airguns (500 cubic inches) comprised the sound source, and the 1,200-m streamer was configured for 48 channels, allowing 12-fold coverage. The record length was 7 s, with a 2 ms sampling rate. The data are high quality and reveal far more structural detail than do the older single channel seismic data¹.

Figure 3 summarizes the new structural interpretation of the Kerguelen Plateau province. The northern sector includes the province's only islands, composed of volcanics and volcanics/intrusives⁸, and offshore comprises a major sedimentary basin⁹ as well as numerous smaller basins separated by igneous intrusions or basement horsts^{1,2,6,7,9–11}. A large positive gravity anomaly coincides with the north-west end of the plateau, defining an apparent structural entity we term the 'Kerguelen massif'. The massif, sub-rectangular except for a narrow ridge extending east-northeast from the northeastern margin, includes the Kerguelen Islands in the south and small sedimentary basins in the north and north-east. South-southeast of the Kerguelen massif, and separated from it by a transverse gravity trough, is an amalgamation of generally positive gravity anomalies extending as far as Heard Island. The largest of the positive anomalies

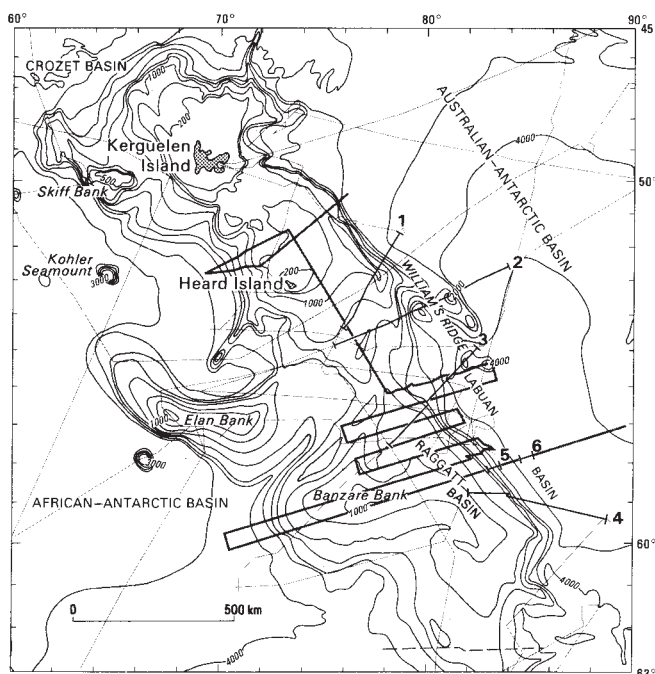
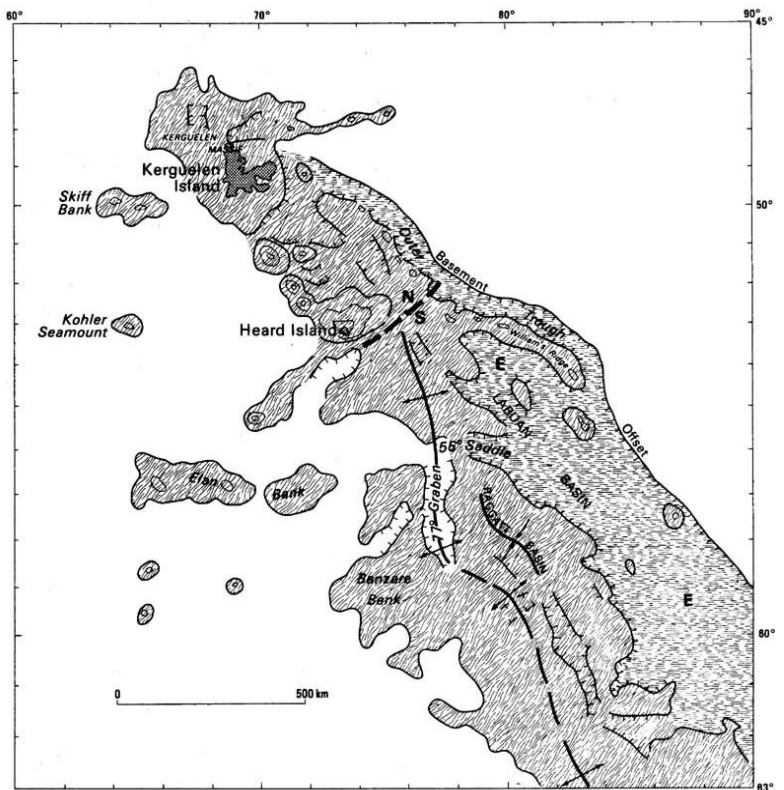


Fig. 1 Bathymetry of the Kerguelen Plateau based on shipboard echo-sounding and the Seasat-derived free air gravity field; contour interval is 500 m. Seismic control is indicated as follows: thick solid lines, 48-channel seismic; dashed lines, digital single-channel seismic; thin solid lines, analogue single-channel seismic. All lines have gravity, magnetics, and bathymetry. Numbered segments (1–6) indicate profiles displayed in Fig. 4.

Fig. 3 Morphotectonic map of the Kerguelen Plateau province based on combined analysis of Seasat-derived gravity and seismic reflection data. Diagonal stippling indicates elevated areas; horizontal stippling indicates deep basins. Scarp symbols indicate moderate to steep gravity gradients, in some cases coinciding with normal faults. N/S, boundary between northern and southern sectors; E, eastern sector. Anticline symbol indicates axis of broad arch of southern sector and syncline symbol indicates axis of Raggatt sedimentary basin. The names William's Ridge and Labuan Basin have been formally proposed to the Australian Hydrographer for these newly-defined features.



extends for some tens of kilometres north-east and north-west from Heard, and probably demarcates the volcanic pedestal of the island. Other discrete, strong sub-circular positive anomalies extend north-west from Heard Island in a rectilinear chain; the major sedimentary basin⁹ lies north-east of the chain. A series of more subdued positive anomalies extends north-west along the north-east margin of the plateau. Where MCS data are available, we observe that the positive anomalies coincide with volcanic pedestals, horsts, or peaks of intrusive rock^{6,7}. Intervening irregularly-shaped negative anomalies coincide with small, well-developed sedimentary basins. The positive anomalies along the north-east margin, although not obviously horsts, lie along trend from horst blocks of the southern and eastern sectors.

The boundary between the northern and southern sectors lies on a transverse line immediately south-east of Heard Island, and is marked by a distinct change in gravity signature and structural style. The bulk of the southern plateau is characterized by a more subdued gravity field and a simpler structural regime than the northern sector. The southern sector comprises a single broad anticlinal arch, disturbed by normal faults, but, unlike the northern sector, not obviously intruded by igneous rocks. Fault motion appears to have occurred over a long time interval^{6,7}. In places faulting, non-deposition and/or current erosion have resulted in exposure of igneous basement on the sea floor. From the boundary with the northern sector at ~53°S southwards to a prominent saddle at ~56°S, the anticlinal arch plunges gently south-southeast; it is bounded to the north-east by a series of horsts and grabens⁹ but the south-west flank is a simple, gentle slope. Coincident with the saddle at 56°S, the arch may be cut by a transverse graben. South-southeast of the saddle the arch rises and reaches maximum development at ~58°S, where it is ~500 km across and rises to ~700 m depth. Banzare Bank forms the southwest flank of the arch at these latitudes. The axis of the arch is marked by a graben which trends almost north-south, the 77° Graben¹; half-grabens on the flanks face towards the central axis. The Raggatt Basin (Fig. 3), and seismic horizons within it, extend over hundreds of kilometres, and sediment thickness in places is >2 s two-way time^{6,7} (tw).

The two major morphotectonic components of the eastern sector are a broad, deep morphological basin (Labuan) and a bounding ridge (William's). These features are characterized by prominent gravity gradients and, in the case of the Labuan Basin, by the broadest negative gravity anomalies of the entire Kerguelen Plateau province. Figure 4a illustrates the apparent continuity of structure from the horsts and grabens of the north-east margin of the northern and southern sectors into the basins and ridges of the eastern sector. Figure 4b displays active boundary faults between the eastern and southern plateau sectors. The faults coincide with a sharp gravity, gradient, and earthquakes have occurred along trend <35 km north of the seismic line^{12,13}. The deep basement of the Labuan Basin extends to the southeasternmost extent of the seismic data (~62°S), and its associated negative gravity signature to the limits of the Seasat-derived gravity data at 63°S. The Labuan Basin seems to extend farther to the Antarctic margin. Figure 4c shows basement in the basin; it is faulted, averages 7–8 s (tw) in depth, reaches maximum depths of ≥8 s (tw), and has a sediment cover of 0–≥2 s (tw) thickness. Discrete sub-circular strong positive gravity anomalies at 55.5°S and 57.5°S are probably seamounts. William's Ridge is ~500 km long. Profile 2 in Fig. 4a is the only complete seismic transect of the feature. The ridge stands ~1,500 m above the adjacent ocean floor, is ~100 km across, and appears to consist of a series of horst blocks. North-east of Labuan Basin and William's Ridge is a sharp gravity gradient which corresponds to changes in basement depth, morphology, and structure^{1,6,7}. North of ~55°S the transition observed on seismic data is abrupt¹, and to the south more gradual^{6,7}. This basement offset (Fig. 3) marks the boundary of the eastern sector of the Kerguelen Plateau province with the Tertiary oceanic crust dated by marine magnetic anomalies².

The recognition of an eastern sector of the Kerguelen Plateau province and marked differences in the structure of the three plateau sectors implies a more complex evolution of the feature than has been proposed^{1,14–16}. Specifically, there is a need to review previous fits of the Kerguelen Plateau province and Broken Ridge. Previous reconstructions^{1,14,15} have employed parameters derived from the separation of Australia and

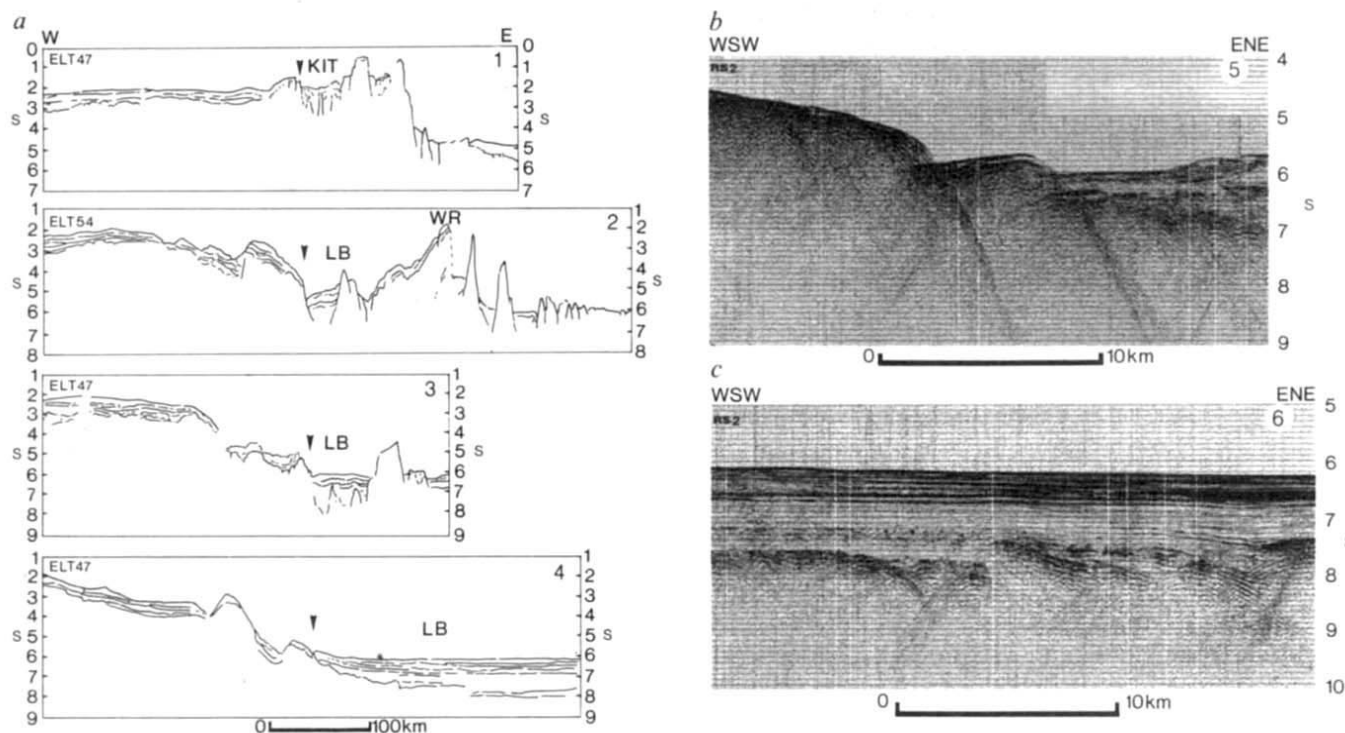


Fig. 4 Seismic data documenting the eastern sector of the Kerguelen Plateau province. ELT, line drawings of *Eltanin* single-channel seismic data; RS, *Rig Seismic* MCS data. Profiles are numbered in upper right hand corner; see Fig. 1. *a*, Sequence of line drawings across the north-east margin of the Kerguelen Plateau province. Grabens of the Kerguelen Island Trough (KIT)¹ form an apparent rift zone high on the plateau (1), descend (2) into the Labuan Basin (LB; 3), and persist far to the south (4). William's Ridge (WR), previously mapped as an isolated seamount on pre-Seasat bathymetric maps^{1,2}, is on the north-east end of profile 2, which is the only existing complete transect of the feature. The drawings are aligned (black arrows) along the western flank of the Kerguelen Island Trough and Labuan Basin. *b*, Seismic line showing detail of major faulting marking the boundary between the southern and eastern plateau sectors. Earthquakes in 1973 and 1981 occurred along trend <35 km to the north¹³. *c*, Seismic line showing basement character of Labuan Basin.

Antarctica; we suggest a new reconstruction employing Seasat-derived fracture zone azimuths between the Kerguelen Plateau and Broken Ridge. Dating of the eastern sector and of the Diamantina Zone, presumed conjugates⁷, is vital to any tectonic reconstruction for the region.

The relationship between the Kerguelen Plateau and hotspot volcanism needs further attention. Flexure studies and/or sampling are needed to test if certain apparently compensated components of Kerguelen Plateau province—Elan Bank, Skiff Bank and Kohler seamount—are volcanic edifices whose origins may be unrelated to those of the generally uncompensated main northern, southern, and eastern plateau sectors. The rectilinear chain of seamounts between Kerguelen and Heard islands suggests a slow north-west movement of the Antarctic plate in the interval 39–10 Myr, contradicting previously determined Antarctic plate motions^{1,17,18}. Resolution of the conflicting evidence on absolute motion of the Antarctic plate awaits further sampling and age dating of the volcanic rocks on the plateau. The southern

plateau sector, in having far greater relief, active major faulting, and no evidence of Neogene volcanism, differs from typical hotspot swells¹⁹. Gaussberg volcano on Antarctica, although chemically distinct from Kerguelen Plateau volcanics²⁰, lies along the trend of the plateau—is it part of the plateau province? Although further geophysical studies may provide clues, Ocean Drilling Program operations on the Kerguelen Plateau planned for 1987–88 have the most potential for resolving questions of the feature's origin and evolution.

We thank the master, crew and the scientific staff of the R/V *Rig Seismic* for their roles in acquiring the multichannel seismic data. C. Brenner, M. Weigel and the BMR Drawing Office provided technical assistance. We thank S. Cande, J. Colwell, P. Hill, M. Munsch, C. Pigram, D. Ramsay, R. Schlich, J. Sclater, H. Stagg and B. Willcox, for critically reading the manuscript. M.F.C. and H.L.D. publish with the permission of the Director, Bureau of Mineral Resources; Lamont-Doherty contribution 4057.

Received 2 January; accepted 28 August 1986.

- Houtz, R. E., Hayes, D. E. & Markl, R. G. *Mar. Geol.* **25**, 95–130 (1977).
- Schlich, R. in *The Ocean Basins and Margins 6: the Indian Ocean* 51–148 (Plenum, New York, 1982).
- Tapley, B. D., Born, G. H. & Parke, M. E. *J. geophys. Res.* **87**, 3179–3188 (1982).
- Haxby, W. F., Karner, G. D., LaBrecque, J. L. & Weissel, J. K. *Eos* **64**, 995–1004 (1983).
- Haxby, W. F. & Weissel, J. K. *J. geophys. Res.* **91**, 3507–3520 (1986).
- Ramsay, D. C. *et al. Bur. Miner. Resour. Aust. Rep.* **270**, (1986).
- Ramsay, D. C. *et al. Geology* **14**, 589–593 (1986).
- Upton, B. J. G. in *The Ocean Basins and Margins 6: the Indian Ocean* 585–648 (Plenum, New York, 1982).
- Wicquart, E. thesis, *Univ. Pierre et Marie Curie, Paris* (1983).
- Schlich, R., Delteil, J., Moulin, J., Patriat, P. & Guillaume, R. C. *r. hebdo Séanc. Acad. Sci., Paris* **272**, 2060–2063 (1971).
- Fröhlich, F. *et al. C. r. hebdo Séanc. Acad. Sci., Paris* **297**, 153–156 (1983).

- Okal, E. A. *Earth planet. Sci. Lett.* **52**, 397–409 (1981).
- Adams, R. D. & Zhang, B. M. *Geophys. J. R. astr. Soc.* **79**, 697–703 (1984).
- Mutter, J. C. & Cande, S. C. *Earth planet. Sci. Lett.* **65**, 369–376 (1983).
- Mutter, J. C., Hegarty, K. A., Cande, S. C. & Weissel, J. K. *Tectonophysics* **114**, 255–279 (1985).
- Mammerickx, J. & Sandwell, D. J. *geophys. Res.* **91**, 1975–1988 (1986).
- Duncan, R. A. *Tectonophysics* **74**, 29–42 (1981).
- Goslin, J. & Patriat, P. *Tectonophysics* **101**, 221–244 (1984).
- Crough, S. T. A. *Rev. Earth planet. Sci.* **11**, 165–193 (1983).
- Tingey, R. J., McDougall, I. & Gleadow, A. J. W. *J. geol. Soc. Austr.* **30**, 241–246 (1983).
- Hayes, D. E. & Vogel, M. *General Bathymetric Chart of the Oceans 5.13 (GEBCO) scale 1:10,000,000* (Canadian Hydrographic Service, Ottawa, 1981).
- Fisher, R. L., Jantsch, M. Z. & Comer, R. L. *General Bathymetric Chart of the Oceans 5.09 (GEBCO) scale 1:10,000,000* (Canadian Hydrographic Service, Ottawa, 1982).

Oxygen isotopes and sea level

J. Chappell* & N. J. Shackleton†

* Biogeography and Geomorphology, Research School of Pacific Studies, Australian National University, Canberra, Australia

† Godwin Laboratory for Quaternary Research, University of Cambridge, Free School Lane, Cambridge CB2 3RS, UK

From the time that detailed oxygen isotope records derived from foraminifera living in the constant-temperature environment of the abyssal ocean became available, there has been a discrepancy between the ice volume record that these records imply, and that derived from the altitude of dated coral terraces around the world. Here, we re-examine the data and conclude that the temperature of the abyssal ocean has been an actively varying component of the climate system.

Oxygen isotope ratios in foraminifera from deep-sea sediments have been used as indicators of past climate following Emiliani's pioneering work¹. $\delta^{18}\text{O}$ in foraminifera varies with $\delta^{18}\text{O}$ in the water from which their carbonate tests are deposited, but differs from the water value by an amount that is temperature dependent². The mean $\delta^{18}\text{O}$ of the whole ocean varies with the quantity of isotopically light ice stored on the continents so that the record from foraminifera in a particular core is a blend of global (ice volume) and local (temperature) components. As the record in planktonic foraminifera may be affected by differences in surface salinity³ as well as by changing seasonal growth and depth habitat⁴, benthic foraminifera have recently been preferred as they provide a less complicated record of global ice volume. It has been assumed that the deep ocean, at least in the Pacific, is so cold that its temperature may be regarded as constant as well as spatially uniform.

Differences exist between $\delta^{18}\text{O}$ records from different cores, even those based on benthic foraminifera. Recent studies have isolated systematic differences between good quality benthic records from different regions which have been interpreted as indicating that the deep Atlantic, which at present is $\sim 2^\circ\text{C}$ warmer than the Pacific, was significantly colder than this during glacial times than at present⁵. However, in these studies it was only possible to study inter-oceanic gradients; a change in the average temperature of the whole deep ocean could not be distinguished from a whole-ocean isotopic change.

We have made a detailed comparison between a deep Pacific $\delta^{18}\text{O}$ record from core V19-30 (ref. 6), and a high-resolution sea-level record from the Huon Peninsula, New Guinea. Although neither can be regarded as a perfect monitor of global ice volume, both records must be dominated by a common ice-volume signal. In making this comparison, we extend earlier discussions which noted a discrepancy between sea-level records of the past 120 kyr as indicated by raised marine terraces, and by the marine oxygen isotope record^{7,8}. The record from core V19-30 has been fully documented as has the methodology on which it is based⁶. The Huon sea-level record⁹ is less well known and its derivation is reviewed to clarify the assumptions involved in comparing it with the $\delta^{18}\text{O}$ record.

Sea-level changes over the past 300 kyr have been interpreted from a flight of raised coral terraces along the tectonically rising north-east coast of Huon Peninsula. Individual reef terraces were dated using the ^{14}C and $^{230}\text{Th}/^{234}\text{U}$ methods, which have given good internal consistency within morphological units¹⁰⁻¹⁴. Each reef developed when the rising sea level overtook the rising land; reef crests represent approximately the peaks of each transgression^{11,13}. Terrace heights increase southeastwards along the coast with increasing uplift rates. The morphology and internal structure of each reef indicates the course of sea-level change relative to the rising land, so that sea level relative to stable ocean floor can be extracted for each section if the uplift rate for that section is known.

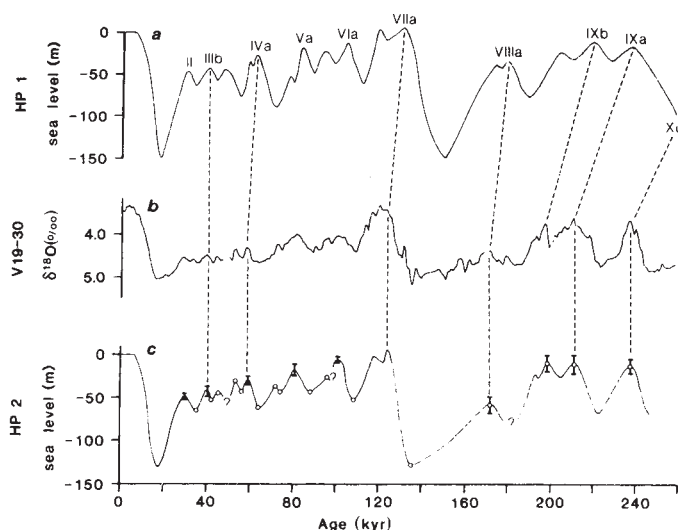


Fig. 1 a, Detailed sea-level curve for Huon Peninsula from ref. 9, with small notches eliminated. b, $\delta^{18}\text{O}$ record for the past 340 kyr from east equatorial core V19-30, from measurements of benthic *Uvigerina senticosus*, plotted to an orbitally-tuned timescale and smoothed with a three-point running mean. c, Detailed sea-level curve for Huon Peninsula as recalculated after detailed correlation with the $\delta^{18}\text{O}$ record from core V19-30.

We have estimated uplift rates assuming that sea level at ~ 125 kyr BP was 6 m above the present level⁹⁻¹². The spread in radiometric analytical data for this high sea-level event has been interpreted as indicating that this episode may have been prolonged. However, stable isotope measurements in both molluscs¹⁵ and corals¹⁶ have confirmed that this interval of high sea level coincided with the rather short substage 5e in the marine oxygen isotope sequence. Reef complex VII represents the '125-kyr marker' in the Huon sequence, and is divided into units VIIa and VIIb^{9-12,17}.

Sea levels associated with other Huon reefs were estimated as follows: given the uplift rate U_j based on the elevation of reef VII in section j , sea level $S_{i,j} = H_{i,j} - U_j t_i$, where $H_{i,j}$ is the elevation of the crest of reef i on section j and t_i is the age of reef i . Sea levels have been estimated for the reef sequence at seven surveyed sections with uplift rates ranging from 0.9 m kyr^{-1} (north-west end of the flight) to 3.5 m kyr^{-1} (south-east end of the flight)¹². Error estimates for the sea level associated with each reef crest are based on results from all sections, combined with the contribution of dating errors⁹. The lower sea levels which intervene between reef crests were estimated in the same way, using the heights of shallow marine and littoral deposits in the Tewai section⁹, supported by dates from other sections^{13,14}.

The detailed sea-level curve for the past 260 kyr derived by this method is shown in Fig. 1a, modified from the previous version only in that small steps of a few metres amplitude have been smoothed out. These steps were interpreted from wave-cut notches throughout the terrace flight; they seem to represent individual seismic events and thus have no global validity⁹. The case for regarding Fig. 1a as a global curve rests on points of agreement with results from elsewhere. Sea levels at 105, 83 and 60 kyr agree closely with estimates from Barbados¹² and Timor¹³ as well as being supported by dates from the Ryukus¹⁹ and in Vanuatu (New Hebrides)²⁰. Sea level was estimated in these other areas using the same method, with the reef or terrace of 125 kyr being the reference surface in each case. Agreement between sites which are widely separated and which have very different uplift rates supports the global use of the Huon curve. However, less support is available for the low points, because

Table 1 Heights, uplift rates and sea-level estimates for Huon Terrace sections

Huon Reef	Age (kyr adjusted)	Tewai	Kanzarua	Blucher	Kwambu	Nama	Sambero	Mean sea level
VIIa	124	440	330	280	220	160	150	+6 (assumed)
Uplift (m kyr ⁻¹)		3.5	2.61	2.21	1.72	1.24	1.16	(from VIII a)
VIa	100	338*	250	215	160	115	110	
		-12	-11	-6	-12	-9	-6	-9 ± 3
VIIb	96	312						
		-24						
Va	81	260	190	155	117	90	80	
		-23	-21	-24	-22	-10	-14	-19 ± 5
Vb	72	216						
		-36						
IVa	59	178	125		70	48		
		-29	-29		-31	-25		-28 ± 3
IVb†	53	156						
		-30						
IIIa	45	112						
		-46						
IIIb	40	98	70	41	28	10		
		-42	-34	-47	-41	-40		-41 ± 4
II	28	52	30	18	7			
		-46	-43	-44	-41			-44 ± 2

Height data are from ref. 9 (Tewai) and ref. 12 (other sections). Heights of reef VIIa at all except Tewai and Kwambu are as given for VIIb in ref. 12. Uplift rate is based on VIIa = 12 kyr, sea level at +6 m.

* Estimated, not surveyed, height. All other heights are from theodolite survey.

† Heights of IVb and IIIa based only on Tewai section, as uncertainty exists over whether IIIa surveyed on other sections is equivalent to 53- or 45-kyr isotope points.

the low sea-level deposits are generally buried by those of the subsequent transgression. Low sea-level points are a good test of the global applicability of the curve because isostatic and other factors will probably have their greatest effect between glacial maxima and minima.

The low sea level at 18 kyr corresponding to the last glacial maximum is an important point for comparison with ¹⁸O records. Figure 1a shows this level at -150 m based mainly on results from northern Australia where lagoonal facies at -135 m (ref. 21), intertidal beach rock at -150 m (ref. 22), and coral below a terrace at -165 m (ref. 23) occur at sites on the outer shelf margins and are dated between 14 and 18 kyr. Lesser values, around -90 to -110 m are reported from the west Atlantic²⁴ and Texas Gulf²⁵. Variations around the globe are to be expected as a result of adjustment to glacial/interglacial changes of ice and ocean volumes²⁶⁻²⁸, although the magnitude of predicted variation differs from one geophysical model to another. The sea-level change between 18 kyr and the present on a broad shelf would be expected to be larger than that recorded on a dipstick-type margin such as the Huon Peninsula²⁸⁻³⁰, so that 150 m may be an overestimate of the 18-kyr level at Huon. Undated evidence does occur at Huon in the form of a sub-merged terrace running along the submarine margin. This is narrow and coralline where it occurs in front of coral terraces, and broad and gravelly where it lies off the flank of a broad deltaic plain¹¹. Correcting its level for local uplift puts the sea level at -130 m for each locality¹¹. In view of this consistency, we use this estimate of -130 m in our subsequent discussion.

The ¹⁸O record from core V19-30, based on *Uvigerina senticosus*, is shown in Fig. 1b. The original sampling interval of 3 cm corresponds to ~500 yr on average although this value varies down the core as accumulation rate varies^{6,31}. Figure 1b shows the record using a timescale that was developed by the SPECMAP project³², and smoothed using a three-point running mean. This timescale was developed iteratively, the starting point being radiometric age determinations both in marine sediments and in coral terraces. The final timescale is developed by tuning the initial record on the basis of its relationship to orbital precession, obliquity and eccentricity functions^{32,33}. Although

the final timescale is not dependent in detail on particular radiometric dates, the modifications are within the uncertainties of the age estimates at all points^{32,33}. In fact, the orbital theory does not provide sufficient constraints on the timescale within stage 3 (in the range 30-60 kyr BP) so that no significant departures from the chronology implied by direct radiometric dating have been used in this interval.

We consider that for the part of the record older than stage 3, the orbitally tuned timescale should be used to refine the age estimates for the Huon Peninsula reefs, on the grounds that independent lines of evidence converge in support of this mechanism, and that the assumption of orbital forcing places constraints on the timescale from 70 kyr back past the useful limits of U-series dating that are much tighter than those imposed by the analytical uncertainties on radiometric dates in this interval. Figure 1c shows the Huon sea-level record as re-estimated on this basis; each sea-level maximum and minimum has been assigned the same age as the assumed correlative in the V19-30 record. The amplitude of some peaks differs significantly between Fig. 1a and c because the sea-level estimate for any point in the sequence depends on the assumed age for that point. The sea level also depends on the value of the uplift rate adopted in each section which depends on the age assumed for reef VII. It also depends on whether this age, and the 6-m calibration height, is assigned to terrace VIIa or VIIb. Finally in traverses containing a single feature for reef VII the uplift rate depends on whether this feature is assumed to represent VIIa or VIIb. Because oxygen isotope studies clearly associate reef VIIa (the older) with substage 5e (ref. 17), and also associate the undifferentiated but dated level in Barbados^{15,16} and in Curacao with 5e, it must be assumed that VIIa is the calibration level to make a comparison consistent with the deep-sea ¹⁸O record. There are various sources of data from the Huon Peninsula⁹⁻¹⁴, so to remove confusion, Table 1 lists reef crest heights from each surveyed traverse together with the age that we have adopted for each, and the resulting sea-level estimate.

Before 130 kyr the sea-level curve derived from marine terraces is subject to larger uncertainties because both the assigned ages and assumed uplift rates become less secure. However, for

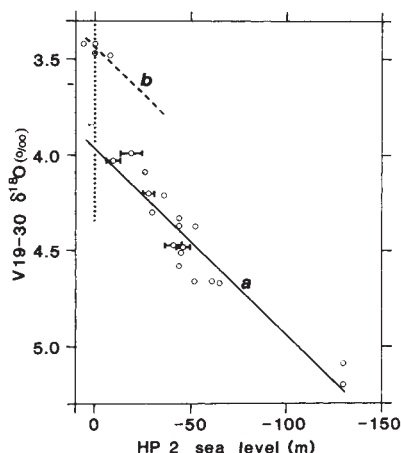


Fig. 2 Huon sea level (recalculated age model, Fig. 1c) versus equivalent ^{18}O in V19-30, from data in Table 2. Regression line *a*, $\delta^{18}\text{O} = 3.96 - 0.0097$ sea level, fitted to all points from 17 to 112 kyr, $r^2 = 0.86$. Line *b* is drawn parallel to the regression line *a*, but passes the area at top left which encompasses data for 5e (124 kyr) and recent (6 kyr-present)

the past 130 kyr we regard Fig. 1c as the best sea-level record available for the New Guinea/Australian region (noting that the magnitude of sea-level oscillations may be smaller in other parts of the world as a result of isostatic factors). Within the past 130 kyr the greatest uncertainty relates to the reef IVb-IIIa area where the isotopic record shows little structure (see Table 2); in Fig. 1c a gap is shown over this part. Table 2 lists sea levels for most turning points in Fig. 1a,c as well as the ^{18}O values for corresponding points in V19-30.

Comparing Fig. 1b with Fig. 1c, it is particularly striking that the sea-level peaks within the last glacial cycle, and especially the well controlled peaks at ~106 and 81 kyr, are relatively closer to the Recent and 124-kyr levels than the equivalent ^{18}O values. Figure 2 plots sea level against ^{18}O for all established turning points in Fig. 1c. A cluster of points around zero sea level and +3.4‰ corresponds to full interglacial stage 1 and substage 5e. A single line cannot be drawn through all data points: that is, there is not a simple linear relationship between sea level and ^{18}O in benthic foraminifera. Although this situation could result in part from temporal variations in the isotopic composition of the stored ice, this cannot be the chief factor involved. If the 0.6‰ isotopic difference between the interglacial points and the two terraces at ~-20 m were to be explained solely in terms of the removal of isotopically light water from the ocean, this would necessitate ice with an ^{18}O content of ~100‰. The lightest ice in Antarctica is ~-60‰^{34,35}. The average isotopic composition of the ice in an ice sheet depends mainly on the isotopic composition of the snow falling in the cold central region, which depends on the air temperature and hence mainly on the altitude of the central part. Thus the first ice accumulating in small Northern hemisphere ice sheets must have been less isotopically light than was the ice stored in the fully developed ice sheet.

One possible solution to this difficulty could be that a significant amount of isotopically light floating ice accumulated³⁶⁻³⁸. This would change the ocean isotopic composition without having any effect on sea level. The only place where a sufficient quantity of floating ice can plausibly be envisaged is the Arctic Ocean; the detailed and continuous stratigraphic sequences that have recently been documented from deep-sea cores in the Eastern Arctic Basin³⁹ exclude the possibility that this ingenious idea is correct, at least for the 18 kyr glacial maximum. The data in Fig. 2 require a mechanism that operated during warm substages 5a and 5c as well as at glacial maximum time.

A more probable explanation is that this isotopic shift results

Table 2 Huon sea levels and V19-30 ^{18}O values

Huon Reef	Age HP1 (kyr)	Sea level HP1 (m)	Age HP2 (kyr)	Sea level HP2 (m)	$\delta^{18}\text{O}$
Modern	0	0	0	0	3.42
I	6	0	6	0	3.47
I/II	18	-150	17	-130	5.09
II	30	-46	29	-46	4.48
II/IIIb	34	-62	35	-65	4.67
IIIb	40	-40	40	-41	4.47
IIIb/a	43	-57	42	-52	4.66
IIIa	47	-42	44	-45	4.51
IIIa/IVb	55	-78	?	?	—
IVb	58	-36	53	-30	4.30
IVb/a	60	-40	56	-44	4.58
IVa	62	-27	59	-28	4.20
IVa/Vb	72	-88	64	-61	4.66
Vb	77	-52	72	-36	4.21
Vb/a	79	-60	74	-44	4.33
Va	83	-18	81	-19	3.99
Va/VIb	89	-48	88	-44	4.37
VIb	95	-22	96	-26	4.09
VIb/a	99	-38			
VIa	105	-12	106	-19	4.03
VIa/VIIb	112	-60	112	-62	4.38
VIIb	120	+4	118	0	3.49
VIIb/a	123	-8	122	-8	—
VIIa	132	+6	124	+6	3.37
VII/VIII	150	-145	135	-130	5.20

from a temperature effect on the isotopic composition of the benthic foraminifera analysed. A best-fit line through all the glacial points in Fig. 2 intercepts the axis at ~+3.95‰ (line *a* on Fig. 2), requiring a temperature difference of ~2°C. The slope of this line is 0.97‰ per 10 m (very close to the generally adopted value⁴⁰) which corresponds to a mean ^{18}O content for the abstracted water of ~-35‰, a reasonable value.

Deviations from this regression are still significant; in Fig. 3 we examine the temporal pattern of the deviations from line *b* on Fig. 2, which parallels line *a* and passes the Recent points, to evaluate the relative contributions of varying sea level, deep-water temperature and ice isotopic composition to the ^{18}O record. For convenience, Fig. 3 has a scale representing deep-water potential temperature at the site of core V19-30 (refs 41, 42). This scale would be valid only if there were no variation in the isotopic composition of stored ice-sheet ice. If this assumption is correct, Pacific deep waters cooled rapidly to ~0°C by ~110 kyr, and then cooled more slowly to ~-1°C around 40 kyr, warming to 0°C at 18 kyr and then to 1.5°C at present. Although a warming at 18 kyr seems surprising, one model has predicted such an effect⁴³. However, an alternative would be that the stored ice (which includes Antarctic ice) became gradually more negative between 110 and 30 kyr (ref. 44), and that the very rapid advance of ice to lower latitudes between 30 and 20 kyr involved more temperate ice that was less isotopically negative. Further work will be needed to establish the possible contribution of temperature variability to the pattern in the glacial section of Fig. 3; we regard the overall glacial cooling of Pacific deep water by 1.5°C that is implied, as inescapable.

It would be useful to compare the older sections of the Huon and V19-30 records, but this is not possible for several reasons: (1) there is no earlier sea-level datum comparable to that at 124 kyr to use for uplift-rate calibration; (2) the errors in $^{230}\text{Th}/^{234}\text{U}$ dates become more serious in this interval; Fig. 1 suggests that this method may actually overestimate ages in the 180-240 kyr age range; (3) only two sections have been surveyed for the pre-5e part of the Huon record. It is regrettable that the morphostratigraphy of this important Upper Quaternary area has been surveyed and described by only one geologist (J.C.).

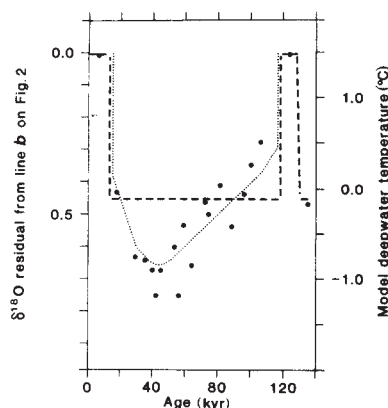


Fig. 3 Residuals from line *b* (Fig. 2) versus estimated age. Scale to right shows calculated potential temperature at the sea floor at the site of core V19-30, if this line is interpreted as the correct relationship between sea level and ocean $\delta^{18}\text{O}$, and deviations from the line ascribed to temperature. Heavy dashed line is our preferred interpretation of the temperature record; the light dotted line is an alternative.

No other area is known where the details of sea-level change are available for field examination.

We conclude that deep waters in the Pacific Ocean were $\sim 1.5^\circ\text{C}$ cooler in glacial and interstadial times than in the short (~ 10 kyr duration) interglacials of substage 5e and the present. One interpretation of our analysis is that deep water was coldest at ~ 40 kyr and that almost 1°C of warming occurred by the glacial maximum and a further 1.5°C of warming subsequently. Alternatively, part of the observed detail resulted from changes in the isotopic composition of the ice sheets, with the deep Pacific remaining $\sim 1.5^\circ\text{C}$ cooler than the present temperature throughout the glacial. Regardless of this detail, massive cooling of the already cold abyssal ocean must be considered an important aspect of the ice age climate.

This work was supported in Cambridge by the NERC. N.J.S. acknowledges stimulating discussion with many colleagues, particularly J.-C. Duplessy.

Received 26 March; accepted 22 August 1986.

- Emiliani, C. *J. Geol.* **63**, 538–538 (1955).
- Epstein, S., Buchsbaum, R., Lowenstam, H. A. & Urey, H. C. *Bull. geol. Soc. Am.* **62**, 417 (1951).
- Epstein, S. & Mayeda, T. *Geochim. cosmochim. Acta* **4**, 213–224 (1953).
- Berger, W. H., Killingley, J. S. & Vincent, E. *Oceanol. Acta* **1**, 221–230 (1978).
- Duplessy, J.-C., Moyes, J. & Pujol, C. *Nature* **286**, 479–481 (1980).
- Shackleton, N. J., Imbrie, J. & Hall, M. A. *Earth planet. Sci. Lett.* **65**, 233–244 (1983).
- Broecker, W. S. *Science* **188**, 1116–1121 (1975).
- Dodge, R. E., Fairbanks, R. G., Benniger, L. K. & Maurrasse, F. *Science* **219**, 1423–1424 (1983).
- Chappell, J. *Search* **14**, 99–101 (1983).
- Veeh, H. H. & Chappell, J. *Science* **167**, 862–865 (1970).
- Chappell, J. *Bull. geol. Soc. Am.* **85**, 553–570 (1974).
- Bloom, A. L., Broecker, W. S., Chappell, J., Matthews, R. K. & Mesolella, K. J. *Quat. Res.* **4**, 185–205 (1974).
- Chappell, J. & Polach, H. A. *Bull. geol. Soc. Am.* **87**, 235–240 (1976).
- Chappell, J. & Veeh, H. H. *Nature* **276**, 602–603 (1978).
- Shackleton, N. J. & Matthews, R. K. *Nature* **268**, 618–620 (1977).
- Fairbanks, R. G. & Matthews, R. K. *Quat. Res.* **10**, 181–196 (1978).
- Aharon, P., Chappell, J. & Compston, W. *Nature* **283**, 649–651 (1980).
- Chappell, J. & Veeh, H. H. *Bull. geol. Soc. Am.* **89**, 356–368 (1978).
- Kohishi, K., Omura, A. & Nakamichi, O. *Proc. 2nd int. Coral Reef Symp.* **2**, 595–613 (1973).
- Neef, G. & Veeh, H. H. *Nature* **269**, 682–683 (1977).
- Van Andel, T. J. & Veevers, J. J. *Bur. Miner. Res., Aust. Bull.* **3** (1967).
- Jongsma, D. *Nature* **228**, 150–151 (1970).
- Veeh, H. H. & Veevers, J. J. *Nature* **226**, 536–537 (1978).
- Milliman, J. D. & Emery, K. O. *Science* **162**, 1121–1122 (1970).
- Curry, J. *Recent Sediments, Northwest Gulf of Mexico* (ed. Sheppard, F. P.) 221–266 (American Association of Petroleum Geologists, Tulsa, 1960).
- Walcott, R. L. *Quat. Res.* **2**, 1–14 (1970).
- Cathles, L. M. in *Earth Rheology, Isostasy and Eustasy* (ed. Morner, N. A.) 11–44 (Wiley, New York, 1980).
- Clark, J. A., Farrell, W. E. & Peltier, W. R. *Quat. Res.* **9**, 265–287 (1978).
- Bloom, A. L. *Bull. geol. Soc. Am.* **78**, 1477–1494 (1967).
- Chappell, J. *Quat. Res.* **4**, 405–428 (1974).

- Shackleton, N. J. & Pisias, N. G. in *Atmospheric CO_2 : Natural Variations, Archaeal to Present* (eds Sundquist, E. T. & Broecker, W. S.) 303–317, Am. geophys. Un. monogr. **32** (1985).
- Imbrie, J. *et al.* in *Milankovich and Climate* (eds Berger, A. *et al.*) 269–305, (Reidel, Dordrecht 1984).
- Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121–1132 (1976).
- Dansgaard, W., Johnson, S. J., Moller, J. & Langway, C. C. *Science* **166**, 377–381 (1969).
- Picciotto, E., Cameron, R., Crozaz, G., Deutsch, S. & Wilgain, S. *J. Glaciol.* **7**, 273–287 (1968).
- Mercer, J. H. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **8**, 19–27 (1970).
- Broecker, W. S. *Science* **188**, 1116 (1975).
- Williams, D. F., Moore, W. S. & Fillon, R. H. *Earth planet. Sci. Lett.* **56**, 157–166 (1981).
- Markussen, B., Zahn, R. & Thiede, J. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **50**, 271–284 (1985).
- Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39–55 (1973).
- O'Neil, J. R., Clayton, R. N. & Mayeda, T. K. *J. chem. Phys.* **51**, 5547–5558 (1969).
- Shackleton, N. J. *Colloq. Int. cent. Nat. Res. Scient.* **219** (1974).
- Newell, R. *Quat. Res.* **4**, 117–126 (1974).
- Mix, A. C. & Ruddiman, W. F. *Quat. Res.* **21**, 1–20 (1984).

Predicting aqueous aluminium concentrations in natural waters

Christopher S. Cronan, William J. Walker* & Paul R. Bloom†

Department of Botany and Plant Pathology, University of Maine, Orono, Maine 04469, USA

* Land and Water Resources Center, University of Maine, Orono, Maine 04469, USA

† Department of Soil Science, University of Minnesota, St. Paul, Minnesota 55108, USA

Aluminium is a pH-sensitive element that can cause acute toxicity symptoms in some organisms at aqueous activities of $10\ \mu\text{M}$ or less^{1–3}. Scientists working on agricultural systems have long been concerned with the deleterious effects of aluminium on crop roots^{4,5}. More recently, environmental scientists have reported a potentially harmful biogeochemical link between acidic deposition onto forest soils and aluminium toxicity in forest and aquatic communities of northeastern North America and northern Europe^{6–8}. Because of this general interest in aluminium toxicity as an environmental threat, there have been renewed efforts to model the chemistry and transport of aqueous aluminium in soils and surface waters. Here we propose that much of the spatial and temporal variability in aqueous aluminium chemistry can be accounted for by a two-component equilibrium model involving a solid-phase humic adsorbent and an aluminium trihydroxide mineral phase. Inputs for the model are solution pH, copper-extractable organic aluminium and the titratable carboxyl content of soil humus.

Several workers^{9,10} have found that $[\text{Al}^{3+}]$ (aquo aluminium ion activity) in surface waters can be predicted accurately as a function of pH using an aluminium trihydroxide solubility relationship. However, other reports^{11–13} have presented data revealing aluminium concentrations that are much lower than predicted—that is, despite pH values less than 4.5 and elevated concentrations of complexing organic ligands, the solutions contained low concentrations of soluble aluminium. This illustrates the failure of existing geochemical models to account for all the major variations in $[\text{Al}^{3+}]$ and we refer to it as the 'paradox of aluminium undersaturation'.

Figure 1 shows $[\text{Al}^{3+}]$ for a variety of natural waters from North America and northern Europe. The expected $p\text{Al}^{3+}$ values for solutions in equilibrium with natural gibbsite and amorphous $\text{Al}(\text{OH})_3$ are also shown. As illustrated, there are three major points where observed and predicted aqueous aluminium activities diverge: the upper soil rooting zone, downstream of wetlands and during stream peak flows. While it has been suggested that these cases of lowered $[\text{Al}^{3+}]$ are the result of kinetic constraints on the dissolution of mineral aluminium¹⁴, there have been no rigorous efforts to confirm the kinetic hypothesis in natural systems. In this study, our objective was to derive a model that could account for these 'undersaturated'

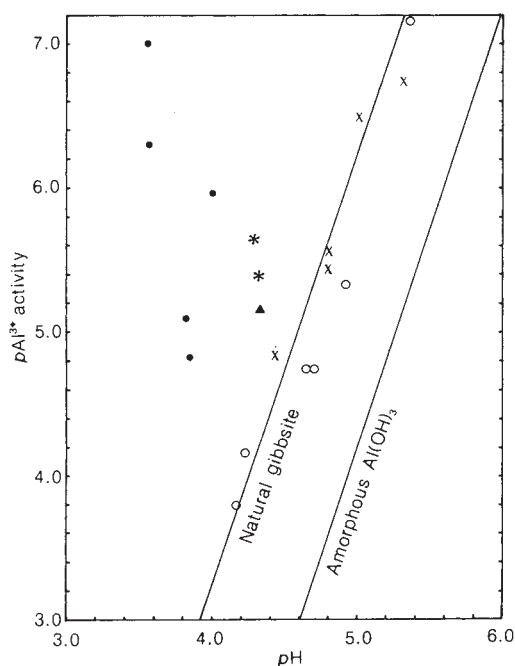


Fig. 1 Aquo aluminium activity as a function of pH for a variety of natural waters from North America and Europe. Field samples include upper O/A horizon soil solutions (●), lower B/C horizon soil solutions (○), wetland drainage (*), stream storm flow (▲), and stream baseflow (×) from the following watersheds: Woods Lake, New York, Huntington Forest, New York, Pancake-Hall Creek, New York, Solling spruce forest, West Germany, Plastic Lake, Ontario, Lange Bramke, West Germany, Camp Branch, Tennessee, Birkenes, Norway and Lake Gardsjon, Sweden (C.S.C. and C. L. Schofield, unpublished data).

waters, so that $[Al^{3+}]$ could be predicted across the range of inorganic and organic environments illustrated in Fig. 1.

Soil and solution chemistry data for this study were obtained using methods described in Table 1. Laboratory experiments were conducted with field-moist, 2 mm sieved forest floor horizons (Oi and Oa)¹⁵ from contrasting northern and southern watersheds in the United States and Canada. Aqueous aluminium ion activities were estimated using the GEOCHEM program¹⁶, appropriate thermodynamic constants¹⁷, and the following mass balance equation:

$$Al^{3+} = MAL - MOAL - Al(OH)_x^{3-x+} - (\text{complexes of } F^- \text{ and } SO_4^{2-})$$

where,

MAL = total monomeric aluminium

MOAL = organic monomeric aluminium

As shown in Fig. 1, many of the data points from natural waters lie between the solubility lines for gibbsite and amorphous $Al(OH)_3$. This general pattern agrees with the results of previous studies^{10,18} and implies that $[Al^{3+}]$ in these waters may be largely controlled by an aluminium trihydroxide solid phase. In order to model solution aluminium chemistry for these types of waters, one can take an average solubility product of $\log K^0 = 8.77$, corresponding to natural gibbsite¹⁸ or interlayer aluminium hydroxide¹⁹. Then $[Al^{3+}]$ in solution can be related to the aluminium hydroxide solid-phase by the following relationship:

$$\log [Al^{3+}] = 8.77 - 3 \text{ pH} \quad (1)$$

Combining this with measured concentrations of complexing ligands and appropriate thermodynamic data¹⁷, it is possible to estimate the equilibrium distribution of inorganic aqueous aluminium.

Table 1 Apparent solubility expressions for predicting Al^{3+} activity in the presence of soil organic matter at different bound aluminium ratios

Bound aluminium ratio	Slope	Intercept	Solubility expression	pAl^{3+} at pH 4.40
0–0.10	0.40	3.62	$pAl = 0.40(pH) + 3.62$	5.38
0.20–0.30	0.60	2.52	$pAl = 0.60(pH) + 2.52$	5.16
0.60–0.70	1.04	0.26	$pAl = 1.04(pH) + 0.26$	4.84
0.90–1.00	1.34	–1.28	$pAl = 1.34(pH) - 1.28$	4.62

Calculated values of pAl^{3+} at pH 4.40 are illustrated for each case. Regressions are based on Oi and Oa samples from a representative Adirondack Spodosol. Organically bound Al in each soil horizon was determined by extraction with 0.5 M $CuCl_2$, followed by ICP-AES analysis of Al (ref. 21). Subsamples of each soil were hydrogen saturated with 1 M HCl ¹¹ and were then titrated under nitrogen to pH 7.0 in 2×10^{-3} M $CaCl_2$ with 0.01 M NaOH. Titratable acidity to pH 7 was used as an operational estimate of RCOOH groups. The bound Al ratio for each native soil horizon was computed as the equivalents of Cu-extractable Al per unit organic matter divided by the equivalents of RCOOH per unit organic matter. Based upon the titration data, subsamples of Oi and Oa forest floor horizons were amended with $Al(NO_3)_3$ to produce a range of bound Al ratios from 0.1 to 1.0. For each soil horizon and bound Al ratio, subsamples were augmented with 2×10^{-3} M $CaCl_2$ and were adjusted with HCl or NaOH to yield a pH range from 3.0 to 5.0. In addition, untreated control subsamples of all native soil horizons were prepared as soil pastes in H_2SO_4 (pH 4.0) augmented with 2×10^{-3} M $CaCl_2$. All soil mixtures were then equilibrated for 48 h, followed by mechanical vacuum extraction through acid washed cellulose pulp. Aqueous monomeric inorganic and organic Al fractions were estimated using ion exchange separation, reaction with 8-hydroxyquinoline at pH 8.3, and spectrophotometric determination at 395 nm²². Inorganic anions were measured with ion chromatography.

Earlier studies have emphasized the potential importance of soil organic matter in controlling aqueous aluminium, and Bloom *et al.*¹¹ suggested that aluminium-hydrogen ion exchange on organic matter controls $[Al^{3+}]$ in acidic surface soils. We propose that a general model for predicting aqueous aluminium chemistry must incorporate this type of organic solid phase reaction to account for the zone of extreme undersaturation illustrated in Fig. 1.

Soil equilibration experiments were used to determine the influence of humic adsorption-desorption reactions on $[Al^{3+}]$ in Oi and Oa organic soil horizons. Results indicated that (1) forest floor saturated pastes reached stable levels of $[Al^{3+}]$ in 5 min or less and (2) at solution pH values between 3.0 and 5.0, plots of pAl^{3+} versus pH were approximately linear for discrete ranges of bound aluminium, and $[Al^{3+}]$ increased with increasing percentages of solid-phase aluminium (Fig. 2). This pattern indicates that equilibrium conditions can be assumed for these soil horizons and that it is possible to construct $[Al^{3+}]$ apparent solubility lines based upon the bound aluminium ratio and pH.

Table 1 shows the best-fit linear regression solubility expressions for the four experimental ranges of bound aluminium ratio shown in Fig. 2. Combining these equations, gives the following general solubility expression for predicting aquo aluminium ion activity in natural waters dominated by humic adsorption-desorption reactions:

$$pAl^{3+} = b(pH) + c \quad (2)$$

where b is the slope $= 1.05 \times (\text{bound Al ratio}) + 0.345$, and c is the intercept $= -(5.47 \times (\text{bound Al ratio})) + 3.879$. This empirical regression equation confirms the strong relationship between $[Al^{3+}]$, solution pH and carboxyl-bound Al, and provides a simple means of predicting $[Al^{3+}]$ in acidic organic soil horizons.

The same data for aluminium solubility can be modelled more formally using the modified mass action expression derived by Langmuir²⁰. Given the reaction, $3H^+ + AlX \rightleftharpoons H_3X + Al^{3+}$, the empirical exchange function would be $K_{ex} = ([Al^{3+}]/[H^+]^3)(H_3X/AlX)^n$, where K_{ex} and n are constants, the brackets denote activities of the ions, and AlX and H_3X are their mole fractions on organic sorbent X. Log-linearization of this expression yields:

$$\log ([Al^{3+}]/[H^+]^3) = \log K_{ex} + n \log (AlX/H_3X) \quad (3)$$

Table 2 Model validation for a variety of North American forest soils

Watershed	Horizon	pH	Cu Al		RCOOH			
			milliequivalents per gram (Ash-free basis)		Bound Al ratio	pAl^{3+} observed	pAl^{3+} calculated RCOO-Al	pAl^{3+} calculated Al(OH) ₃
Pancake-Hall, New York	Oi	5.00	0.06	1.77	0.04	5.37	5.60*	6.23
	Oa	4.06	0.09	3.13	0.03	5.30	5.25*	3.41
	Bh	4.35	5.52	5.88	0.94	4.30	4.53*	4.28†
	Stream	4.80	—	—	—	5.55	—	5.63†
Huntington, New York	Oi	5.21	0.10	2.10	0.05	5.70	5.70*	6.86
	Oa	4.11	0.80	2.40	0.34	4.51	4.90*	3.56
	Bh	4.41	12.17	12.64	0.96	4.87	4.59*	4.46†
Plastic Lake, Ontario	Oi	5.32	0.03	1.06	0.03	5.81	5.72*	7.19
	Oa	4.55	0.09	2.02	0.04	5.39	5.42*	4.88
Woods Lake, New York	Oi	4.40	0.07	1.75	0.04	5.59	5.37*	4.43
	Oa	4.02	1.11	5.20	0.21	4.86	4.99*	3.29
	Bh	4.09	1.52	1.62	0.93	4.57	4.20*	3.50
Camp Branch, Tennessee	O	5.20	0.17	2.17	0.08	5.99	5.67*	6.83
	Stream	5.02	—	—	—	6.49	—	6.29†
Univ. Forest, Missouri	O	5.30	0.04	0.84	0.05	6.54	5.71	7.13

Observed values for pAl^{3+} were obtained with soil paste extracts, while calculated pAl^{3+} values were derived with either the humic-aluminium or aluminium trihydroxide model components. The former model (equation 2) was calibrated with experimental data from the Pancake-Hall Adirondack watershed; the apparent solubility lines from that Spodosol were broadly representative of the other soils.

* Close agreement (± 0.4) between pAl^{3+} observed and humic-Al model.

† Close agreement (± 0.4) between pAl^{3+} observed and inorganic solubility model.

Table 2 contains a summary of the model validation results obtained for a number of soils from North America with bound aluminium ratios ranging from 0.03 to 1.00. Using the actual bound aluminium ratio and solution pH for each soil sample, the humic-aluminium partition model was used to predict solution $[Al^{3+}]$ for the untreated native soil horizons from each watershed. These predictions were compared with observed soil paste $[Al^{3+}]$ values and with predictions from the $Al(OH)_3$ solubility model. The fit of a model was judged satisfactory if it correctly predicted pAl^{3+} to within ± 0.4 units. The humic-aluminium model was successful in predicting the pAl^{3+} in all the O horizons with the exception of the sample from the Missouri watershed. Whereas the $Al(OH)_3$ solubility model provided an accurate prediction for pAl^{3+} in stream water at low flow, the model was unable to predict pAl^{3+} correctly in any of the O horizons.

We examined the region of the stability field where the humic-aluminium and $Al(OH)_3$ apparent solubility lines intersect. The Oi horizon data in Table 2 shows that surface organic horizons with higher pH (>4.8 to 5.0) and low mineral content will often follow the humic-aluminium solubility line and can actually maintain oversaturation with respect to the solubility of interlayer $Al(OH)_3$. This is probably due to the scarcity of clay nucleation sites in these horizons and interference of soluble organic matter in the formation of mineral precipitates. However, as the pH is raised above about 5.2, it is likely that amorphous $Al(OH)_3$ forms by homogeneous nucleation. Thus, the upper limit of applicability for the humic-aluminium model is probably about pH 5.2.

In practice, this model will typically apply to surface organic horizons and peaty hydrologic source areas, whereas the inorganic solubility model will be appropriate for lower mineral soil horizons and for the groundwater zone. During peak flows in lower order streams, it is likely that stream water chemistry will be strongly influenced by 'undersaturated' lateral flow derived from organic-rich surface soil horizons. Under these conditions, where streamwater results from mixing of lateral flow and base-flow, it will be necessary to combine accurate hydrologic flux estimates with these chemical models.

The model presented here can account for much of the variation in $[Al^{3+}]$ observed in natural waters draining different inorganic and organic environments. Although we previously suspected that waters undersaturated in $[Al^{3+}]$ could result from

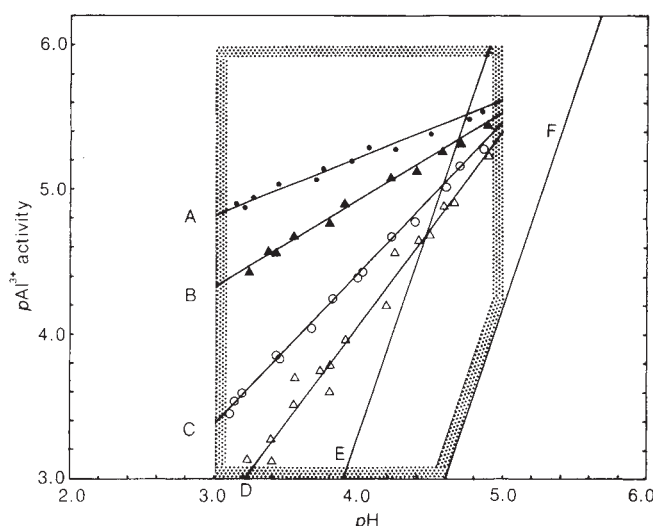


Fig. 2 Stability field for apparent aluminium solubility in the presence of soil humic material in the pH range 3.0 to 5.0. The points along each apparent solubility line represent equilibrium pAl^{3+} values for Oi and Oa horizons from an Adirondack Spodosol. These horizons were pretreated to yield the following four different ranges of bound aluminium ratio: line A = 0 to 0.1; line B = 0.2 to 0.3; line C = 0.6 to 0.7; and line D = 0.9 to 1.0. The solubility lines for interlayer (E) and amorphous $Al(OH)_3$ (F) at 25 °C are shown for reference.

kinetic constraints on Al dissolution and desorption, our results suggest that these waters may actually be in equilibrium with a solid phase humic adsorbent. With further testing and development, this model may be suitable for application to studies of aluminium transport associated with acid deposition.

We thank C. L. Schofield for contributing to many of the Al speciation measurements that provoked this investigation and B. Lee for typing the manuscript. This is a contribution from the ALBIOS project, supported by contract RP 2365 from the Electric Power Research Institute.

Received 13 January; accepted 6 August 1986

1. Haug, A. *CRC Crit. Rev. Pl. Sci.* **1**, 345-373 (1984).
2. Baker, J. P. & Schofield, C. L. *Wat. Air, Soil Pollut.* **18**, 289-309 (1982).

3. Hue, N. V., Craddock, G. R. & Adams, F. *Soil Sci. Soc. Am. J.* **50**, 28-34 (1986).
4. Hardy, F. *J. agric. Sci.* **16**, 616-630 (1926).
5. Pavan, M. A. & Bingham, F. T. *Soil Sci. Soc. Am. J.* **46**, 993-997 (1982).
6. Cronan, C. S. & Schofield, C. L. *Science* **204**, 304-306 (1979).
7. Ulrich, B., Mayer, R. & Khanna, P. K. *Soil Sci.* **130**, 193-199 (1980).
8. Henriksen, A., Skogheim, O. K. & Rosseland, B. O. *Vatten* **2**, 255-260 (1984).
9. Christophersen, N., Seip, H. M. & Wright, R. F. *Wat. Resour. Res.* **18**, 977-996 (1982).
10. Nordstrom, D. K. & Ball, J. W. *Science* **232**, 54-56 (1986).
11. Bloom, P. R., McBride, M. B. & Weaver, R. M. *Soil Sci. Soc. Am. J.* **43**, 488-493 (1979).
12. Cronan, C. S. *Oikos* **34**, 272-281 (1980).
13. Driscoll, C. T., Van Breemen, N. & Mulder, J. *Soil Sci. Soc. Am. J.* **49**, 437-444 (1985).
14. Hooper, R. P. & Shoemaker, C. A. *Science* **229**, 463-465 (1985).
15. Soil Survey Staff *Soil Taxonomy*. (Soil Conservation Service, USDA Washington D.C., 1981).
16. Mattigod, S. V. & Sposito, G. *Am. chem. Soc. Symp. Ser.* **93**, 837-856 (1979).
17. Nordstrom, D. K., Valentine, S. D., Ball, J. W., Plummer, L. N. & Jones, B. F. *U.S.G.S. Water-Resources Investigation Report* 84-4186 (1984).
18. Johnson, N. M., Driscoll, C. T., Eaton, J. S., Likens, G. E. & McDowell, W. H. *Geochim. cosmochim. Acta* **45**, 1421-1437 (1981).
19. Bloom, P. R., McBride, M. B. & Chadbourne, B. *Soil Sci. Soc. Am. J.* **41**, 1068-1073 (1977).
20. Langmuir, D. in *Adsorption from Aqueous Solutions* (ed. Tewari, P. H.) 1-18 (Plenum, New York, 1981).
21. Hargrove, W. L. & Thomas, G. W. *Soil Sci. Soc. Am. J.* **48**, 1458-1460 (1984).
22. Driscoll, C. T. *Int. J. envir. analyt. Chem.* **16**, 267-283 (1984).

A new Miocene hominoid from Kenya

R. E. Leakey & M. G. Leakey

National Museums of Kenya, PO Box 40658, Nairobi, Kenya

A newly discovered early Miocene site in northern Kenya has yielded a partial cranium, several mandibles, isolated teeth and some associated postcranial elements of a large ape that is distinct from other known African and Asian hominoids. The new genus implies a greater complexity in the relationships between the Miocene hominoids and may offer useful evidence concerning the origin of the Asiatic and African large apes. The genus *Sivapithecus* has been reported from East Africa¹ but the new material supports the hypothesis that this genus only occurs beyond Africa², although having an African ancestry. The new genus presents a mosaic of characters which suggest an early radiation of Miocene hominoids.

The specimens were found in exposed sediments along the western side of Lake Turkana in northern Kenya in 1985-86 at a small but rich early Miocene locality named Kalodirr (3°20' N, 35°40' E). Approximately 10 km south there are other early Miocene fossil-bearing sediments known by several names, including Losodok³, Moruaret Hill⁴, and Moruarot^{3,5,6}, which probably belong to the same sequence of Miocene deposits. It has not yet been possible to date the locality but we estimate the age may be between 16 and 18 million years based on faunal comparison. Almost all the elements of the mammalian macrofauna recovered are represented at Rusinga (Hiwegi Formation)⁷ and the fauna is similar to that from Buluk¹ and Moruarot⁶ (Table 1), but three hominoid taxa were recovered all representing new genera. The largest is described here.

Order Primates
Superfamily Hominoidea
Family incertae sedis
Afropithecus gen. nov.

Generic diagnosis. A large hominoid with long narrow palate, distinguished from all known large hominoids by the long distinct muzzle and steeply inclined frontal such that in lateral profile, prosthion, rhinion, glabella and bregma are on almost the same line, and by the corresponding obliquity of the mandibular ramus. It differs from *Sivapithecus*, *Ramapithecus*, *Kenya-pithecus*, *Proconsul* and *Rangwapithecus* in the lingual and buccal basal flare of the upper molars and premolars, particularly P⁴, so that the cusp tips tend to be relatively close compared to the cervical width. It differs from *Sivapithecus*, *Ramapithecus* and *Kenya-pithecus* in the buccolingually expanded premolars, the pronounced cingulum bordering the protocone of the upper molars mesially and lingually and the buccolingually broad lower incisors, particularly I₂. It differs from *Rangwapithecus* in the buccolingually expanded upper

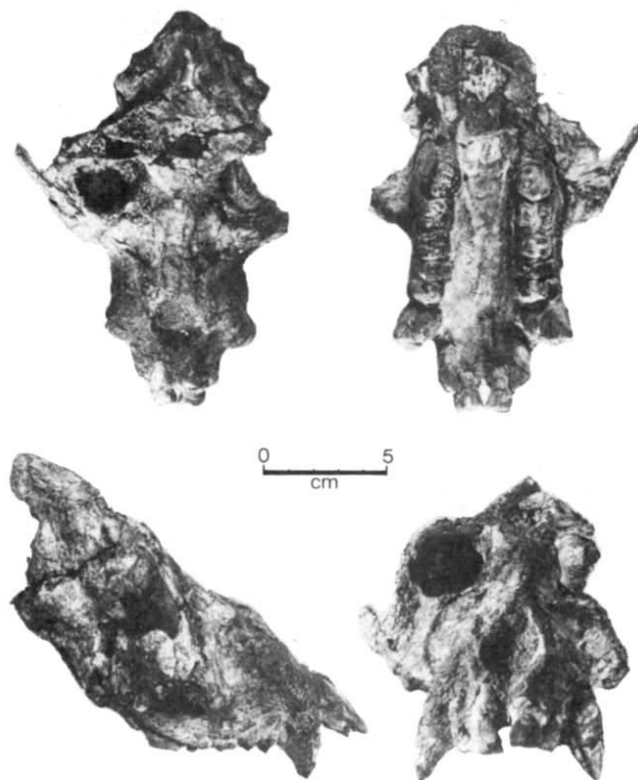


Fig. 1 *Afropithecus turkanensis* cranium, KNM-WK 16999 (type). Above, superior and inferior aspects; below, right lateral and frontal aspects.

premolars, the upper P⁴ which is shorter mesiodistally than P³, the less developed lingual cingulum of the upper molars, the procumbent upper incisors and the long superior aspect of the premaxilla. It differs from the similar sized *Proconsul major* in mandibular morphology particularly the lack of development of the superior transverse torus and the steeply sloping sub-incisive planum.

Etymology. *Afro* after the continent on which it is found, *pithecos* Greek for ape.

Afropithecus turkanensis sp. nov.

Species diagnosis. As for genus

Type specimen. KNM-WK 16999, facial and frontal region of a cranium with dentition (Fig. 1).

Hypodigm.

KNM-WK 16840	lt. mandible
KNM-WK 16901	lower lt. C, root rt. C, rt. fibula, prox. rt. m/t III, rt. m/t IV, dist. m/t I
KNM-WK 16959	weathered rt. and lt. mandible frags.
KNM-WK 16962	lt. maxilla and premaxilla, rt. C, rt. max. frag., rt. frontal frag.
KNM-WK 17010	associated lower teeth
KNM-WK 17011	rt. mandible frag.
KNM-WK 17012	associated upper teeth
KNM-WK 17013	rt. upper I ¹
KNM-WK 17014	mandible frags.
KNM-WK 17016	associated lower teeth, partial shaft rt. ulna
KNM-WK 17021	rt. mandible and associated teeth
KNM-WK 17023	frag. rt. lower C crown
KNM-WK 17024	rt. lower M ₂
KNM-WK 17025	lt. lower M ₁ frag.
KNM-WK 17026	frag. lower I crown
KNM-WK 17037	lt. lower P ₃ crown
KNM-WK 17162	lower I ₁ and rt. P ₃ frag.
KNM-WK 17168	frag. lower I crown

Etymology. After Lake Turkana.

Table 1 Faunal list

Mollusca	Mammalia—continued
<i>Melanoides tuberculata</i>	<i>Paraphymys stromeri</i> †
<i>Limicolaria</i> sp.	<i>Paraphymys</i> sp.
<i>Lanistes carinatus</i>	<i>Elmerimys woodi</i> †
<i>Pila ovata</i>	Cricetodontidae indet.
<i>Etheria elliptica</i>	<i>Kenyalogomys minor</i> †
Insecta	<i>Anasinopa leakeyi</i> †‡
cf. Agromyzidae	<i>Hecubides macrodon</i> †
Pisces	<i>Kichechia zamanae</i> †‡
<i>Myliobatiformes</i> sp. indet	<i>Hyaenodon</i> sp.
<i>Polypterus</i> sp.	<i>Platybelodon</i> sp.*
cf. Bagridae	<i>Prodeinotherium hobleyi</i> *†
Siluriformes indet.	<i>Megalohyrax championi</i> *†‡
Serpentes indet.	cf. <i>Aceratherium acutirostratum</i> *†‡
Chelonina	<i>Masritherium aequitoralis</i> †
Trionychidae indet.	cf. <i>Hyoboops africanus</i> *†
Testudinidae indet.	<i>Kenyasus rusingensis</i> †‡
Crocodylia nov.	<i>Libycochoerus jeanelli</i> †‡
Mammalia	<i>Lopholistriodon moruoroti</i> *†‡
<i>Afropithecus turkanensis</i> *	<i>Diamanthohyus africanus</i> *†‡
Hominoidea gen. nov. A	<i>Canthumeryx sirtensis</i> *†
Hominoidea gen. nov. B	Giraffidae nov.
<i>Pronasilio</i> sp.	<i>Dorcatherium parvum</i> †‡
cf. <i>Gymnurechinus</i> sp.	<i>Dorcatherium pigotti</i> *†‡
<i>Paraphymys pigotti</i> †‡	<i>Dorcatherium chappuisi</i> *†‡

* Buluk (ref. 6).

† Rusinga, Hiwegi Formation (ref. 5).

‡ Moruorot (ref. 4).

Description. The type specimen, KNM-WK 16999, includes the snout, the facial skeleton, the frontal reaching close to the coronal suture, the pterygoid processes, the sphenoid body, the vomer and the relatively unworn adult dentition.

The palate is shallow, long and narrow and the tooth rows converge posteriorly. The internal nasal aperture is remarkably narrow (15 mm) and high (~18 mm). There is a large (6.5 mm) diastema between the strongly procumbent I^2 and the canine. The enamel surface of the incisors and canines is markedly wrinkled by numerous small vertically running grooves and the occlusal surface of the cheek teeth is covered by numerous wrinkles. The slightly worn central incisors are large and broad while the lateral incisors are relatively small and mediolaterally compressed. The canines are broad and in cross section at the cervix almost circular. Their relative size suggests that the individual was male. The premolars are large, P^3 being larger than P^4 . In both the buccal cusp is higher than the lingual and there is marked buccal basal flare. The P^3 has slight lingual basal flare whereas in P^4 this is more marked and the cusp tips are relatively close compared to the basal buccolingual width. $M^1 < M^2 < M^3$ and the molars are broader than they are long (Table 2). On all three molars there is a distinct but short mesial and lingual cingulum extending around the protocone. The enamel cap appears relatively thick judged from the degree of wear before dentine exposure.

The surface bone of the premaxilla is largely replaced by calcite. Its superior aspect is long, forming a gently sloping shelf with a central longitudinal groove. The nasals are long and narrow (13 mm across the widest point) but the tips are lost close to the margin of the small oval pyriform aperture. The left side of the maxilla is best preserved. There is a distinct jugum and posterior to it a small postcanine fossa. On either side of the nasals, anterior to the marked lacrimal foramina, the maxilla is distinctly hollowed. The zygomatic process is deep below the orbit (37 mm) and the inferior border swings markedly upwards from the masseteric tubercle towards the temporal, a striking feature probably related to an unusually high position of the cranial vault. The right orbit is virtually complete and relatively undistorted. It is higher than it is wide and has an asymmetrical outline; the greatest depth (30 mm) is lateral to its longitudinal axis and the greatest width (25 mm) is close to the superior margin. The supraorbital tori are very thin and there is a small but distinct supraorbital notch. The interorbital distance is large

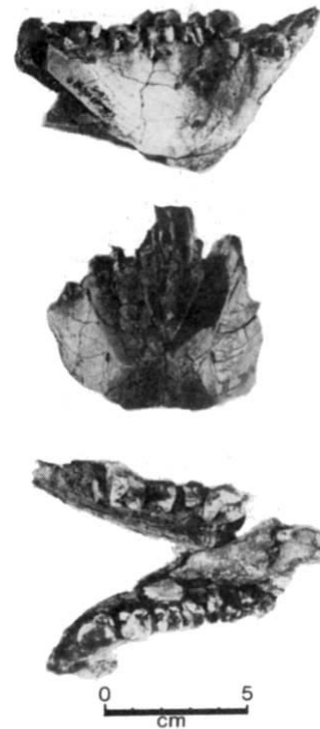


Fig. 2 *Afropithecus turkanensis* mandible, KNM-WK 16840. Above, right lateral aspect; centre, anterior aspect and below, occlusal aspect.

(24.5 mm). The marked temporal lines swing sharply towards the midline forming a frontal trigon which rises steeply towards the apex, approximately 52 mm posterior to glabella. There is a distinct, mostly midline frontal sinus, seen through the breaks in the region immediately above glabella. Glabella does not project anterior to the orbits and is weakly developed. In lateral profile, prosthion, rhinion, glabella and bregma are almost in the same line, a feature not observed in any other fossil or living large hominoid.

KNM-WK 16840 (Fig. 2) is the most complete of the mandibular specimens. The body is deep and robust and the tooth row relatively short. The symphysis is deep (60 mm) and the internal surface slopes steeply down to the genial fossa with little development of the transverse tori. The superior anterior aspect is the widest because of the lateral flare of the canines. The narrowest point is at a level with the mental foramen and below this, it widens towards the prominent mental tubercles. The mandibular body is deepest anteriorly (54 mm at the posterior canine alveolar margin) and becomes slightly shallower towards the posterior (44 mm below $M2$). There is a distinct mandibular fossa. There is a clearly defined and very large digastric fossa. One specimen, KNM-WK 17021, preserves details of the ramus which is set at an oblique angle to the body. There is a slight inferior projection of the mandibular angle.

Several specimens provide details of the lower dentition at various stages of wear (Table 2). The lower incisors are all marked by varying degrees of mesial curvature from the root tip to the crown and the roots are broad labiolingually, particularly I_2 . The P_3 s are relatively large, rather narrow and sectorial with a slight extension of the enamel onto the buccal face of the mesial root. The P_4 s are small compared to the P_3 s, low crowned and bicuspid. The molars are slightly inflated buccally and on one specimen, KNM-WK 17024 there is evidence of a slight buccal cingulum. M_3 has a small tuberculum sextum. $M_1 < M_2 < M_3$.

Afropithecus displays characters typical of a variety of Miocene hominoids combined in a single taxon that should also include the large hominoid recently described from Buluk¹. We

Table 2 *Afropithecus* tooth dimensions (mm)

Tooth	Acc. no.	Upper		Acc. no.	Lower	
		mesiodistal	buccolingual		mesiodistal	buccolingual
I1	16999	11.6	9.2	16840	(5.2)	8.0
				17016*	(4.5)	8.4
				17021*	5.1	8.6
				17162	8.1	5.5
				17010	(8.0)	10.0
I2	16999	7.6	9.7	16840	6.5	9.5
				17016*	6.0	10.5
				17021	6.0	10.8
				16901	12.2	15.4
				17014*	—	(15.0)
C	16999	18.0	17.4	17016	12.0	16.8
				17016	8.9	(14.3)
P3	16999	9.0	14.6	16840	(15.0)	—
P4	16999 17012	7.8 6.5	14.9 12.0	17010	8.0	7.8
				17014*	(12.5)	8.3
				17016*	8.0	(10.0)
				17021	(8.5)	9.6
				17016	(10.3)	—
M1	16999 17012	(11.5) 9.0	(13.0) 11.2	17021	(10.5)	—
				17010	12.5	(10.5)
M2	16999 17012	(12.0) 10.0	(14.0) 11.5	17014	12.0	11.0
				17016	12.4	11.5
				17021	(13.0)	—
				17024*	12.0	10.5
				17010*	14.3	13.0
M3	16999 17012	(12.5) 10.4	(13.0) 10.5	17014*	14.0	—
				17016	14.0	11.0
				17021*	(14.5)	(11.5)
				—	—	—
C-M3	16999	69	—	—	—	—
P3-M3	16999	52	—	—	—	—
M1-M3	16999	36	—	—	—	—

Left tooth dimensions are given except where * indicates right tooth dimensions.

now believe that *Sivapithecus*^{8,9} is restricted to Asia including the Siwaliks, Indo-Pakistan². A careful comparison of the mandibular morphology of *Proconsul major* as represented by the material known from Songhor and Koru¹⁰ differs from *Afropithecus* in the strong development of a superior transverse torus and lack of an inferior torus, the long shallow alveolar subincisive planum and the marked increase in size of the lower molars from M₁ to M₃, not seen in *Afropithecus*. The maxillary morphology of *P. major* is poorly represented. The only reasonably complete specimen that may represent this taxon is the Moroto palate and associated facial fragments. Its assignment¹⁰⁻¹³ to *P. major* is disputed^{14,15} but when the Moroto palate is compared to KNM-WK 16999, the superior aspect of the premaxilla is shorter, the breadth of the nasal aperture is greater, P₄ is less buccolingually expanded and there is no well developed lingual cingulum on the molars.

One weathered maxillary fragment from Moruorot, KNM-LS 7, previously referred to *P. major*¹⁰ was recently designated the type of *Proconsul (Xenopithecus) hamiltoni*¹⁶. The two eroded and worn molar crowns^{17,18} are both elongated buccolingually whereas those of *Afropithecus* are more square. Together with the limited details of cusp morphology, this suggests that this specimen is not *Afropithecus*. In fact the specimen is so poor that we suggest that the name *P. (Xenopithecus) hamiltoni* be abandoned.

Afropithecus has been compared with the poorly represented African genus *Kenyaipithecus*. Little is known of the maxillary morphology of either the later species, *K. wickeri*, or the earlier species, *K. africanus*, but both differ from *Afropithecus* in the dentition, particularly the lack of buccal basal flare on the molars and premolars, the premolars that are not buccolingually expanded and the molars that lack a marked cingulum bordering the protocone. Similarities are apparent in the root of the zygomatic which leaves the maxilla above M₁ and the postcanine fossa of *K. wickeri*. The mandible of *K. wickeri* is known but differs from *Afropithecus* in having a relatively shallow body with a shallow rather gently sloping subincisive planum.

The phyletic status of the large hominoids from the Miocene

of Europe is not agreed¹⁹⁻²¹ but we have considered the published descriptions and made direct comparisons between *Afropithecus* and casts available to us of the European fossils and we are satisfied that *Afropithecus* is quite distinct from the European species of *Dryopithecus*²², *Ouranopithecus*^{21,23}, *Rudapithecus* and *Bodvapiithecus*²².

Afropithecus resembles the specimen GSP 15000 from the Siwaliks²⁴ attributed to *Sivapithecus sivalensis* (formally *S. indicus*²⁰) in the postcanine fossa, the large procumbent central incisors, the relatively small asymmetrical lateral incisors, the long superior aspect of the premaxilla, the small suprorbital tori and the high hafting of the cranial vault onto the face. It differs in the distinct snout and the extraordinary profile with the almost linear alignment of prosthion, rhinion and glabella. These differences also apply when a comparison is made with the known large hominoids from Lufeng²⁵. The Lufeng specimens resemble *Afropithecus*, however, in the conformation of the temporal lines and the wide interorbital distance. Details of sub-nasal morphology which have been used as a basis for discriminating between various hominoids^{26,27} cannot be discerned on KNM-WK 16999.

In conclusion, the mosaic of characters represented in the new genus, which precludes its inclusion in any of the Miocene hominoid families previously defined, suggests a greater complexity in the early Miocene than was previously apparent. The absence of *Afropithecus* from the richly fossiliferous localities of western Kenya where fossil primates are common may indicate ecological or temporal factors. Further work at Kalodirr and other Miocene sites in northern Kenya will investigate these questions.

We thank the Government of Kenya and the Governors of the National Museums of Kenya. The research was financed by the National Museums of Kenya and contributions from The Maytag Family Foundation, Mr and Mrs L. Newitt and the National Geographic Society of Washington, D.C. We also thank Bw Kamoya Kimeu and his prospecting team, particularly Bw Wambua Mangao who found the type specimen, Bw Joseph Mataba who prepared the majority of the specimens, Dr L.

Jacobs, P. Kat and Ms K. Stewart who identified the microfauna, molluscs and fish respectively, and Dr Alan Walker who assisted in many ways.

Received 25 June; accepted 15 September 1986.

1. Leakey, R. E. F. & Walker, A. *Nature* **318**, 173–175 (1985).
2. Kelley, J. & Pilbeam, D. *Am. J. phys. Anthropol.* **66**, 188 (1985).
3. Arambourg, C. *Ann. Paleont.* **22**, 121–146 (1933).
4. Deraniyagala, P. E. P. *Spolia zeylan.* **25**, 1–42 (1948).
5. Walsh, J. & Dodson, R. G. *Geol. Surv. Kenya* **82**, 1–42 (1969).
6. Madden, C. T. *Paleobios* **14**, 1–12 (1972).
7. Pickford, M. *Kenya Palaeont. Gaz.* **1** (National Museums of Kenya, Nairobi, 1984).
8. Pilbeam, D., Smith, R. *Mem. Geol. Surv. Pakistan* **11**, 1–13 (1981).
9. Pilbeam, D. R., Rose, M. D., Badgley, C. & Lipschutz, B. *Postilla* **181**, 1–94 (Peabody Mus. Nat. Hist., Yale University, 1980).
10. Andrews, P. J. *Bull. Br. Mus. (nat. Hist.) Geol.* **30**, 85–244 (1978).
11. Simons, E. L., Andrews, P. & Pilbeam, D. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) 120–146 (Harvard University Press, Cambridge, 1978).
12. Allbrook, D. & Bishop, W. W. *Nature* **197**, 1187–1190 (1963).
13. Pilbeam, D. *Bull. Peabody Mus. nat. Hist.* **31**, 1–185 (1969).
14. Leakey, L. S. B. in *Classification and Human Evolution* (ed. Washburn, S. L.) 32–49 (Viking Fund Publications in Anthropology, No. 37, 1963).
15. Martin, L. J. *Hum. Evol.* **10**, 139–150 (1981).
16. Madden, C. T. *Primates* **21**, 241–252 (1980).
17. Hopwood, A. T. *Ann. Mag. nat. Hist. (Ser. 10)* **11**, 96–98 (1933).
18. Le Gros Clark, W. E. & Leakey, L. S. B. *Fossil Mammals of Africa* (Br. Mus. Nat. Hist.) **1**, 1–117 (1951).
19. Kay, R. F. & Simons, E. L. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R. L. & Corruccini, R. S.) 577–624 (Plenum, New York, 1983).
20. Bonis, L. de & Melentis, J. *C. r. heb. Séanc. Acad. Sci., Paris (D)* **284**, 1393–1396 (1977).
21. Simons, E. L. & Pilbeam, D. R. *Folia primat.* **3**, 81–152 (1965).
22. Kretzoi, M. *Nature* **257**, 578–581 (1975).
23. Martin, L. & Andrews, P. *Cour. Forsch. Inst. Senckenberg.* 6925–6940 (1984).
24. Pilbeam, D. *Nature* **295**, 232–234 (1982).
25. Wu Rukang *Cour. Forsch. Inst. Senckenberg* **69**, 41–48 (1984).
26. Ward, S. C., Kimbel, W. H. & Pilbeam, D. *Am. J. phys. Anthropol.* **61**, 157–171 (1983).
27. Ward, S. C. & Pilbeam, D. R. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R. L. & Corruccini, R. S.) 211–238 (Plenum, New York, 1983).

Table 1 *Turkanapithecus* tooth dimensions (mm)

	Acc. no.	Tooth	mesiodistal	buccolingual
Upper	KNM-WK 16950	C*	10.0	7.9
		P ³	6.8	8.0
	KNM-WK 16957	P ³	5.9	7.7
		P ⁴	5.5	—
	KNM-WK 16950	M ¹	7.0	7.5
		M ^{2*}	8.6	9.2
		M ³	8.5	8.9
		C – M ^{3*}	44.0	—
Lower	KNM-WK 16950	P ³ – M ³	35.2	—
		M ¹ – M ³	24.0	—
		M ₂	9.2	7.0
		M ₃ *	9.2	7.5

Left tooth dimensions are given except where * denotes right tooth dimensions.

troconid. It is distinguished from the larger *Rangwapithecus gordonii* by the upper M³ approximately the same size or smaller than M² and the upper P⁴ smaller than P³. It is distinguished from the slightly larger *Proconsul africanus* by the distinct snout, the wide interorbital distance, the relatively thick supraorbital tori, the broad nasals, the relatively broad and low mandibular ramus and the upper premolars and molars not buccolingually expanded. It is distinguished from the similar sized *Pliopithecus vindobonensis* by the distinct snout, the long nasals, and the zygomatic process of the maxilla approximately vertical with no protruding inferior orbital margin.

Etymology. After Lake Turkana and *pithecus*, Greek for ape.

Turkanapithecus kalakolensis sp. nov.

Species diagnosis. As for genus.

Type specimen. KNM-WK 16950A and B, a partial cranium and mandible (Figs 1 and 2).

Hypodigm.

KNM-WK 16950I-X postcranial elements probably associated with the type specimen: right femur, proximal and distal left radius, left metacarpal I and fragments of phalanges, rib and vertebrae

KNM-WK 16957 fragment of maxilla.

Etymology. After the River Kalakol of which the Kalodirr is a tributary.

Description. The type cranium, KNM-WK 16950, which is slightly smaller than the female *Proconsul africanus* skull (KNM-RU 7290) is reasonably complete and includes the facial skeleton, partial calvaria and examples of all the permanent upper teeth except the incisors. The associated mandible preserves much of the left ramus, the mandibular body and the left M₂ and right M₃. A complete right femur, fragments of proximal and distal left radius, ribs, vertebrae and foot bones are most probably associated with the cranium and mandible but there is a slight element of uncertainty concerning this association, so they are not designated as part of the type material.

The palate is rather shallow and narrow with the tooth rows converging towards the back. A crack, 2.2 mm wide, runs along the centre of the palate making it appear broader than it was. Judged from the position of the back of the palate and the length of the remaining right pterygoid, the internal nasal aperture was high and narrow. The teeth (Table 1) are moderately to heavily worn. The right upper canine is relatively large suggesting the individual was male. The right P⁴ is missing and was lost in life causing some malocclusion; the wear on the right P³ is greater than on either the left P³ or the left P⁴. P³ is significantly larger than P⁴, both are wide buccolingually. The upper molars are rather square in outline but wider than they are long. M¹ < M³ < M². The morphology of the M's is obscured by wear, dentine being exposed over almost the entire occlusal surface. The M²s are distinctive, having marked, beaded mesial and lingual cingula with an unusual accessory cuspule between them. The M³s

A second new Miocene hominoid from Kenya

R. E. Leakey & M. G. Leakey

National Museums of Kenya, PO Box 40658, Nairobi, Kenya

A recently discovered early Miocene site in northern Kenya (see preceding paper¹) has yielded a number of important fossil hominoid specimens including a partial cranium, mandible and postcranial remains of a medium-sized ape that is distinct from all previously reported genera. The relationship of this taxon to other known hominoids is not clear. A new genus of large hominoid was found at the same site¹. The recovery of two new genera of Miocene hominoids which are apparently not represented at the well known and rich Miocene localities of western Kenya may indicate that a different ecology is being sampled.

The specimens were found in 1985–6 at a new Miocene locality named Kalodirr on the western shores of Lake Turkana. The site has not been dated but a faunal comparison suggests it may be between 16 and 18 million years¹.

Order Primates
Superfamily Hominoidea
Family incertae sedis
Turkanapithecus gen. nov.

Generic diagnosis. A shortfaced small hominoid approximately the size of a *Colobus polykomos*. It is distinguished from all known hominoids by the upper M²s, which have an additional cuspule between the pronounced and beaded mesial and lingual cingula, and the upper molars and P⁴ which have small additional cuspules associated with a small buccal cingulum bordering the paracone. It is distinguished from the Oreopithecidae by the relatively low relief of the dental occlusal surface, upper molars not mesiodistally elongated and lower molars not long and narrow, without buccolingual waisting and lacking a cen-

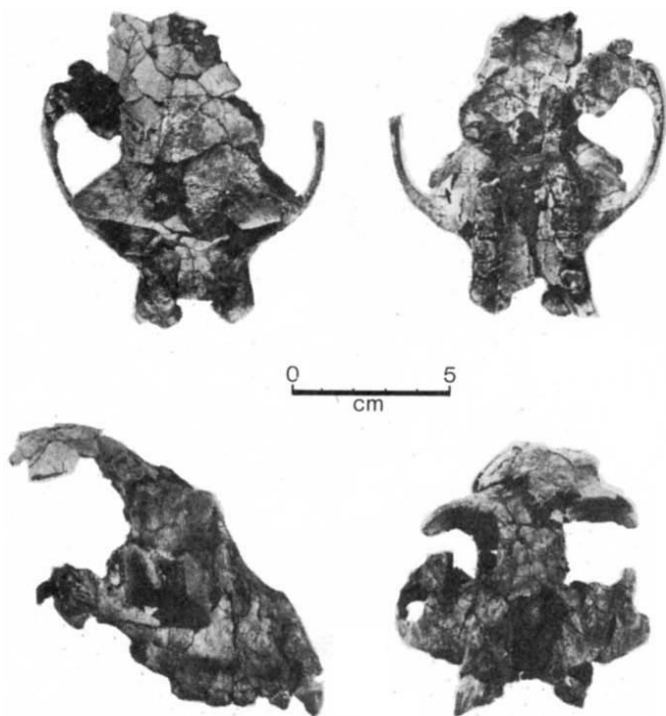


Fig. 1 *Turkanapithecus kalakolensis* cranium, KNM-WK 16950A (type). Above, superior and inferior aspects; below, right lateral and frontal aspects.

are more worn than the M^2 s, probably due to the apparent high set of the lower M^3 s. P^4 to M^3 have small buccal cingula bordering the paracone each with an additional cuspule variably developed.

The premaxilla is largely missing. The pyriform aperture was of moderate size and oval (approx. 17 mm wide). The nasal bones are large, being long (maximum length 27 mm) but also fairly broad (maximum width 14 mm across both). The right maxilla is the best preserved; there is a slight canine jugum but no postcanine fossa. The lower border of the root of the zygomatic process is about 7 mm above M^1 . The zygomatic process is relatively deep below the orbit with the lower border being close to vertical beneath the lower orbital margin. The orbits are incomplete and slightly distorted; they were probably almost circular but slightly higher (approx 24 mm) than wide (approx 22 mm). There is only the slightest supraorbital notch and no development of a supraorbital sulcus. The interorbital distance is large (21 mm). In the region of glabella the bone is crushed and fractured exposing a fairly well developed frontal sinus. The frontal is distorted on the left side lateral to the rather indistinct temporal line. The right temporal line is strongly developed passing initially towards the midline almost parallel with the orbital margin before curving sharply towards the posterior. Little of the parietals is preserved. The occipital area is damaged but there is a fragment that includes the postero-lateral margins of the foramen magnum with most of both condyles and a small portion of bone extending to the nuchal crest. The temporal bone is partially preserved on the right and includes the mandibular fossa, the postglenoid process and the zygomatic process which attaches to the zygomatic.

The mandible is relatively complete particularly on the left side which lacks only some bone in the regions of the angle of the mandible, the coronoid process and the canine alveolus. The mandible is gracile and deeper at the symphysis (28 mm) than below M^2 (21 mm). The internal surface of the symphysis slopes evenly down as a narrow, slightly concave subincisive planum. There is virtually no development of a superior or inferior transverse torus. A small digastric fossa is clearly defined on the right. The anterior root of the broad ascending ramus



Fig. 2 *Turkanapithecus kalakolensis* mandible, KNM-WK 16950B (type). Above, left lateral and below, occlusal aspects.

originates close to M^2 . The mandibular condyle is relatively large and rounded being broad mediolaterally and wide anteroposteriorly. Only two lower teeth were recovered, the moderately worn left M^2 and the right M^3 . The lingual cusps are positioned slightly distal to the buccal and the talonid basin is large. On M^2 , both the mesial and distal foveae are clearly defined and the hypoconulid was probably large.

The phylogenetic relationship of *Turkanapithecus* to other known fossil hominoids is not clear. The cranial morphology has been compared with the only other similar sized African Miocene hominoid for which a skull is known, *Proconsul africanus*^{2,3} and there are important differences. *Turkanapithecus* is distinguished by the distinct and rather 'square' maxilla, the larger pyriform aperture, the longer nasals which widen at either end, the greater depth of the zygomatic process of the maxilla beneath the orbit, the root of the zygomatic process above M^1 not M^2 , the wider interorbital distance, the distinct supraorbital tori and the rather flat supraglabella region. The mandibular ramus is lower and wider anteroposteriorly and the condyle is knoblike and not elongated mediolaterally as in *P. africanus*. Comparison with the similar sized cranium of *Pliopithecus vindobonensis* from the Middle Miocene of Europe^{4,5} also shows differences in the 'square' maxilla of *Turkanapithecus*, the long nasals and the zygomatic process of the maxilla roughly vertical with no protruding inferior orbital margin. The dentition differs from both *Pliopithecus* and *Proconsul* in the features mentioned in the diagnosis.

The dentition has been compared with some isolated teeth from Maboko which have recently been attributed to the previously monospecific family, the Oreopithecidae⁶ and included in a new genus to which *Rangwapithecus vancouveri* is also referred⁷. They are considered to have some affinities with *Rangwapithecus gordonii*. These teeth may be readily distinguished from *Turkanapithecus* in the high cusped elongated molars. Szalay and Delson⁸ suggest that *Oreopithecus* has cercopithecoid affinities on the basis of a number of shared derived characters of the dentition. These features are those which distinguish *Oreopithecus* from *Turkanapithecus*.

Unfortunately little is known of the cranial morphology of the Miocene small bodied apes. *Limnopithecus*, *Dendropithecus* and *Micropithecus*^{9,10} but the dentition clearly separates these genera from *Turkanapithecus*. Although ecological differences may be a factor, it is puzzling, if our age estimates are correct, that *Turkanapithecus* has not been recorded previously from any of the known Miocene sites in western Kenya where fossil apes are common. This relatively complete material from Kalodirr is important in the new evidence it provides of greater diversity in the Miocene Hominoidea.

We thank the Government of Kenya and the Governors of the National Museums of Kenya. The research was financed by

the National Museums of Kenya and contributions from The Maytag Family Foundation, Mr and Mrs L. Newitt and the National Geographic Society of Washington, D.C. We thank Bw. Kamoya Kimeu and his prospecting team who discovered all the fossils described and particularly Bw. Wambua Mangao who found the type specimen. Bw. Joseph Mataba prepared the specimens and Dr Alan Walker assisted in many ways.

Received 25 June; accepted 15 September 1986.

1. Leakey, R. E. F. & Leakey, M. G. *Nature* **324**, 143–146 (1986).
2. Le Gros Clark, W. E. & Leakey, L. S. B. *Fossil Mammals of Africa* (Br. Mus. Nat. Hist.) **1**, 1–117 (1951).
3. Walker, A., Falk, D., Smith, S. & Pickford, M. *Nature* **305**, 525–527 (1983).
4. Hurler, J. A. *Paleont.* **40**, 5–63 (1954).
5. Zapfe, H. *Mem. Suisses Paleont.* **78**, 5–293 (1961).
6. Harrison, T. *Am. J. phys. Anthropol.* **66**, 180 (1985).
7. Harrison, T. *Am. J. phys. Anthropol.* (in the press).
8. Szalay, F. S. and Delson, E. *Evolutionary History of the Primates* (Academic, New York, 1979).
9. Harrison, T. thesis, London Univ. (1982).
10. Fleagle, J. and Simons, E. *Am. J. phys. Anthropol.* **49**, 427–440 (1978).

Vegetation, climatic and floral changes at the Cretaceous-Tertiary boundary

Jack A. Wolfe & Garland R. Upchurch, Jr

US Geological Survey, MS-919, Box 25046, Federal Center, Denver, Colorado 80225, USA

The western interior of North America has the only known non-marine sections that contain the iridium-rich clay interpreted as the Cretaceous-Tertiary (K-T) boundary^{1–7}. Because vegetation and climate can be directly inferred from physiognomy of leaves^{8–15} and because leaf species typically represent low taxonomic categories, studies of leaf floras in these sections provide data on the effects of a terminal Cretaceous event on the land flora, vegetation and climate. A previous study based on detailed sampling of leaves and their dispersed cuticle¹⁶ in the Raton Basin provides a framework for interpretation of other leaf sequences over 20 degrees of latitude. We conclude that at the boundary there were: (1) high levels of extinction in the south and low levels in the north; (2) major ecological disruption followed by long-term vegetational changes that mimicked normal ecological succession; (3) a major increase in precipitation; and (4) a brief, low-temperature excursion, which supports models of an 'impact winter'.

Latest Cretaceous (late Maestrichtian, equivalent to Lancian; Phase 1 (ref. 16)) vegetation from palaeolatitude¹⁷ 46° to >56° N (Fig. 1) was dominantly broad-leaved evergreen (Table 1). In the south small leaf size, paucity of drip tips and thick cuticles with cutinized hairs on both surfaces indicate dry, open-canopy forest^{16,18}. In the Raton sequence (Fig. 2) dispersed-cuticle assemblages from the base of the formation to <1 cm below the K-T boundary show no physiognomic changes that indicate significant change in climate. Although leaf size increased into higher latitudes, drip tips were uncommon, suggesting subhumid forest. At 66° N, vegetation was large-leaved deciduous forest that resulted from the polar light regime¹⁹.

Broad-leaved deciduous plants were not diverse in mid-latitude Lancian vegetation (Fig. 2), as is true for modern Southern Hemisphere vegetation; thus we use the Southern Hemisphere ratio between the proportion of entire-margined species (EM) and mean annual temperature¹⁰. The 20 °C isotherm (megathermal/mesothermal boundary) falls at 50°–53° N, consistent with major floristic differences in megaflores¹⁵ and palynoflores²⁰. The latitudinal temperature gradient was ~0.3 °C per degree of latitude (Fig. 3), a gradient also obtained for the North American Campanian¹⁸. Comparisons of early to late Lancian assemblages in the same areas indicate no significant change in EM and thus no change in temperature.

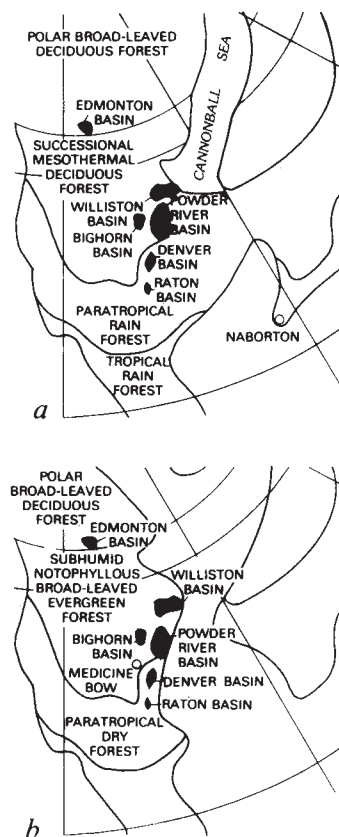


Fig. 1 Maps of the western interior of North America during *a*, the early Palaeocene and *b*, the latest Cretaceous showing major depositional basins, floral assemblages and inferred vegetation. Mesothermal vegetation includes the southern part of polar broad-leaved deciduous forest, subhumid notophyllous broad-leaved evergreen forest and successional mesothermal deciduous forest. Megathermal vegetation includes paratropical dry forest, paratropical rain forest and tropical rain forest (see refs 10, 19). Throughout the Late Cretaceous and early Tertiary the southern margin of polar broad-leaved deciduous forest fell between paleolatitudes 65° and 70° N (refs 18, 19). Continental position adapted from ref. 17; mid-Maestrichtian and early Palaeocene epicontinental seas adapted from ref. 26.

Palaeocene assemblages <3 m above the K-T boundary (Phases 2 and 3) indicate ecological disruption and mass kill. Assemblages of Phase 2, which includes the palynological fern spike⁵, are dominated by herbaceous monocots and/or ferns, which persist into Phase 3. Phase 3 assemblages are low diversity and have leaves of early successional morphology¹². Included are narrow-leaved taxa that typically occurred in disturbed, fluvial habitats during the Lancian, but during early succession following the K-T boundary opportunistically invaded swamp habitats. Temperature inferences are difficult for such a limited flora, but the large leaf size suggests abundant moisture.

All early Palaeocene assemblages of Phases 4 and 5 have large leaf size with no latitudinal increase. Increase in leaf size over Lancian assemblages from 46° to 56° N indicates increased precipitation, as does an increase in the proportion of drip-tipped species and a change in cuticular physiognomy in the south. In the Raton sequence this change occurs at the K-T boundary¹⁶. This increase would lead to high-biomass vegetation and high water tables and was possibly the major causative factor in the widespread initiation of peat deposition that marks the boundary in the northern High Plains. Persistence of high precipitation indicates that the climate was pushed into a new, metastable mode. Initiation of megathermal rain forest environments had major biotic consequences^{18,21}.

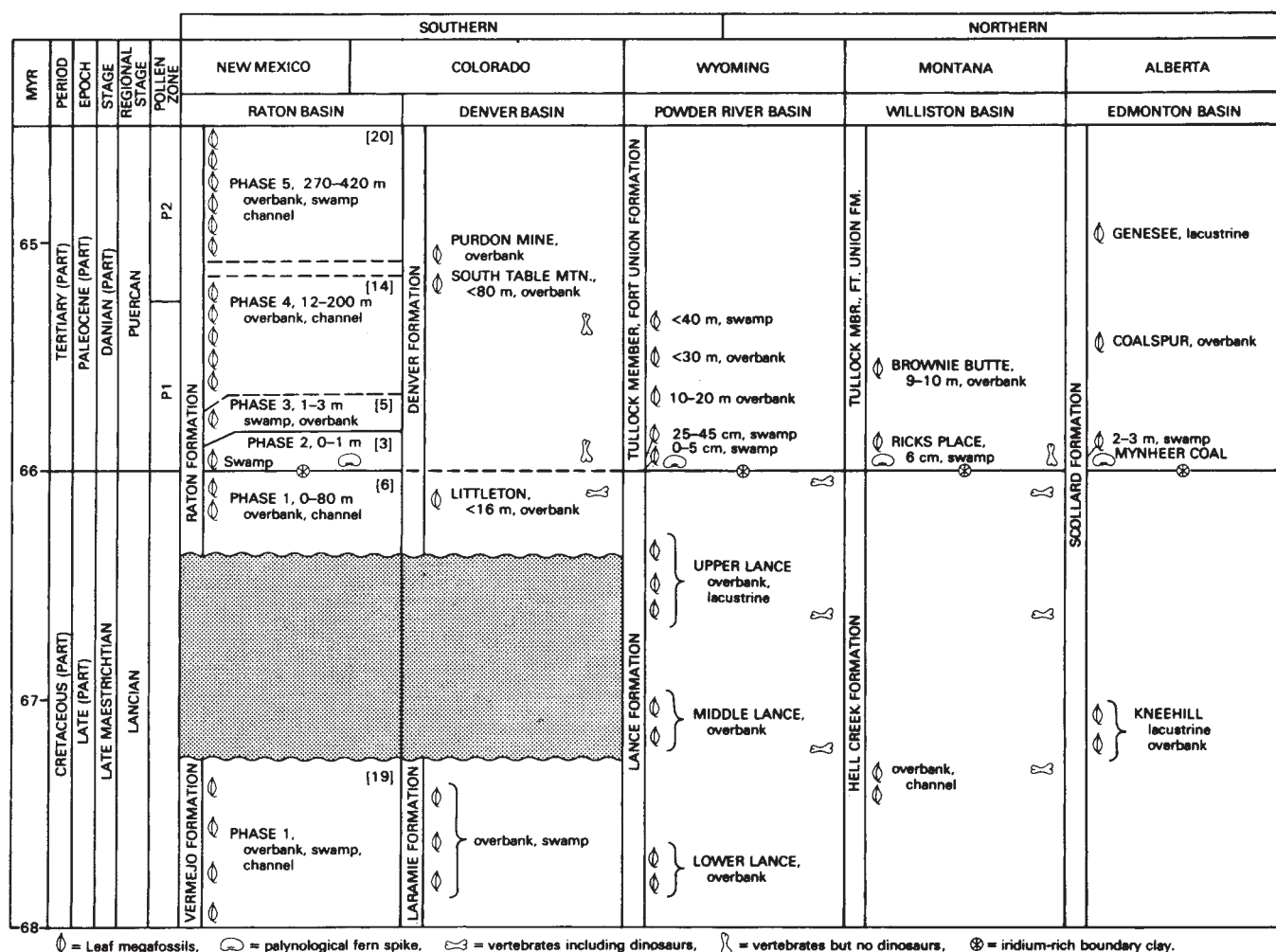


Fig. 2 Stratigraphical and depositional occurrences of some plant megafossil assemblages relative to the Cretaceous-Tertiary boundary in the western interior. Estimates of absolute ages are based on some radiometric and some palaeomagnetic data^{34,35}. Palynological data from refs 5, 36-38; stratigraphical data from refs 16, 39 and unpublished data; Physiognomic and floristic data from refs 16, 40-44 and unpublished data. Following some floral symbols are measured or estimated stratigraphical interval from the boundary. Numbers in brackets in the Raton Basin column are numbers of individual localities.

Diversity in Phases 4 and 5 increased over Phases 2 and 3 but was still lower than in Phase 1. Percentages of entire-margined and evergreen species sharply decline northward compared to the Lancian. During the early Palaeocene, megathermal forest extended to 50°-53° N (as in the Lancian) in the western interior but was anomalously bounded on the north by broad-leaved deciduous forest; mesothermal, notophyllous, evergreen forest is unknown. Anomalous deciduousness characterizes most extant mesothermal vegetation of southeastern North America, resulting from prolonged, intense freezes¹⁰. If such climates characterized the northern High Plains early Palaeocene there would not be a diverse fauna of large ectotherms.

However, occurrence in the northern High Plains of large and presumably ectothermic reptiles (turtles, champsosaurs, and crocodilians²², some >3 m long; W. S. Bartels, personal communication) indicates that earliest Palaeocene temperatures were, as in the Lancian, mesothermal. The Raton data also indicate the same temperature during the Lancian as during the early Palaeocene, and the megathermal/mesothermal boundary was unchanged. These data make previous temperature estimates^{23,24} for the northern high plains early Paleocene, based on deciduousness and low EM, suspect.

Thin leaves, mostly from deciduous plants, typically have toothed margins¹². Thus if vegetation is anomalously deciduous, a low EM relative to temperature results. This phenomenon is

apparent in the western interior early Palaeocene: if the EMs for the leaf assemblages give a valid measure of temperature, the latitudinal temperature gradient from New Mexico to Montana would be 1.2 °C per degree of latitude (Fig. 3). This gradient, which is one-third higher than the present gradient from the southern high plains to Alberta, would produce a mean annual temperature of -2 °C in central Alberta and a polar temperature of -28 °C. Such low temperatures are clearly contradicted by the climatic implications of the Genesee and other northern assemblages²⁵ and would be unlikely because of the transgressing early Palaeocene Cannonball Sea²⁶. Deciduous immigrants in the Denver and elimination of most broad-leaved evergreens in northern Phase 5 assemblages skew the leaf-margin analysis. In Raton Phase 5 vegetation northern deciduous taxa were few and did not affect EM¹⁶.

Oxygen-isotope data²⁷ suggest a temperature increase from the early into the late Maestrichtian, and analyses of North American Maestrichtian floras¹⁸ parallel marine data. No climatic change can be detected between early and late Lancian assemblages in the late Maestrichtian. Lancian temperature is indistinguishable from early Palaeocene temperature.

Palynological interpretations of the K-T boundary in the western interior suggest little extinction, with most taxa persisting in refugia⁵, but are complicated by the facts that (1) palynofloras rarely contain Magnoliidae pollen; and (2) a palynomorph

Table 1 Foliar physiognomy, inferred vegetation and climate and significant taxa of Cretaceous-Tertiary boundary assemblages in the western interior

Area	Assemblage	Phase	Number of species	Number of species	Dicotyledons				Inferred vegetation†	Inferred climate‡	Significant taxa
					entire margin (EM) (%)	Leaf-size index* (%)	Coriaceous (evergreen) (%)	Evergreen drip tipped (%)			
Central Alberta (Edmonton Basin)	Genesee	4-5	26	21	19	73	5	100	PBLDF	W Ms	<i>Metasequoia</i> , trochodendroids, platanoid, <i>Eucommia</i> , <i>Acer</i> , <i>Averrhoites</i> , Hamameloidae, Tiliaceae, Iodeae, Nyctaginaceae
	Coalspur	3-4	3	2	(0)	75	0	—	SPBLDF	W Ms?	<i>Metasequoia</i> , trochodendroid, Hamameloidae
	Mynheer coal (above)	3	8	3	(33)	50	0	—	S	W Ms?	<i>Metasequoia</i> , <i>Glyptostrobus</i> , <i>Thuites</i> , trochodendroid, flacourt, <i>Paranymphaea</i>
	Kneehills, Brazeau	1	21	11	(27)	78	0	—	PBLDF	M Ms	Ginkgophytes, <i>Glyptostrobus</i> , <i>Metasequoia</i> , <i>Nilssonia</i> , trochodendroids, Hamameloidae, <i>Fagopsis</i> , 'Acer', <i>Thuites</i> , flacourt
Eastern Montana (Williston Basin)	Brownie Butte	3-4	15	11	(27)	73	0	0	SMDF	W Ms	<i>Metasequoia</i> , <i>Glyptostrobus</i> , <i>Thuites</i> , <i>Paranymphaea</i> , trochodendroid, platanoid, Hamameloidae, Tiliaceae
	Ricks Place	2-3	4	1	(100)	50	0	0	S	W Ms	aff. <i>Stenochlaena</i> , aff. <i>Averrhoites</i> , herbaceous monocot, leafy bryophyte
	Hell Creek	1	27	21	62	53	76	19	NBLEF	S Ms	Cycad, lauraleans, flacourt, 'Dryophyllum', Hamameloidae, <i>Ginkgo</i> , trochodendroid
Eastern Wyoming (Powder River Basin)	Lower Tullock near Newcastle	3-4	22	12	(33)	—	19	100	SMDF	W Ms	Ferns, <i>Metasequoia</i> , <i>Glyptostrobus</i> , <i>Thuites</i> , <i>Paranymphaea</i> , aff. <i>Averrhoites</i> , trochodendroid, 'Myrtophyllum', 'Carya'
	Lower Tullock near Lance Creek & Glenrock	3-4	17	14	(43)	—	36	60	SMDF/SPRF	W Ms/W Mg	Cycad, <i>Paranymphaea</i> , laurel, <i>Corylites</i> , 'Carya', 'Myrtophyllum', platanoid, trochodendroid, menisperm, palm, Hamameloidae
	Lower Tullock (+25 cm)	3	5	4	(75)	75	75	33	S	W Ms/W Mg	aff. 'Cissites' <i>panduratus</i> , aff. <i>Averrhoites</i> , celastacean, aff. <i>Stenochlaena</i>
	Lowermost Tullock	2	1	0	—	—	—	—	S	?	Herbaceous monocot
	Upper Lance	1	33	24	54	36	76	4	NBLEF	S Ms	Araucarian, lauraleans, 'Dryophyllum', flacourt, euphorb, 'Myrtophyllum', trochodendroid, palm
Southern Wyoming	Middle Lance	1	20	17	(56)	39	82	12	NBLEF	S Ms	'Myrtophyllum', trochodendroid, palm
	Lower Lance	1	36	28	57	35	79	0	NBLEF	S Ms	
	Medicine Bow	1	56	42	67	55	71	13	NBLEF/PDF	S Ms/S Mg	<i>Geinitzia</i> , lauraleans, 'Dryophyllum', Rhamnaceae, 'Myrtophyllum', trochodendroid
Northeastern Colorado (Denver Basin)	South Table Mount. & Purdon Mine	4-5	56	45	62	74	74	44	SPRF	W Mg	Lauraleans, Rhamnaceae, euphorb, 'Myrtophyllum', aff. <i>Grewia</i> , Hamameloidae, palm
	Littleton	1	54	50	71	57	86	10	PDF	S Mg	Diverse lauraleans, diverse euphorbs Rhamnaceae, palm [under study]
	Laramie	1	60	55	71	52	82	11	PDF	S Mg	Diverse lauraleans, euphorbs, <i>Geinitzia</i> , 'Myrtophyllum', palm [needs restudy]
Southeastern Colorado and Northeastern New Mexico (Raton Basin)	Phase 5	5	50	45	74	68	85	55	PRF/SPRF	W Mg	Lauraleans, 'Magnolia', euphorb, aff. <i>Heliocarpus</i> , aff. <i>Grewia</i> , 'Carya', palms 'Magnolia', 'Cissites', celastacean, lauralean, platanoid, <i>Averrhoites</i> , palm
	Phase 4	4	26	25	71	72	90	42	SPRF	W Mg	'Cissites', celastacean, lauralean, ferns, palm
	Phase 3	3	8	5	(80)	70	100	40	S	W Mg	aff. <i>Stenochlaena</i> , leafy byrophyte
	Phase 2	2	0	—	—	—	—	—	S	W Mg	
	Phase 1 (Raton)	1	47	43	72	34	89	9	PDF	S Mg	Diverse lauraleans, diverse euphorbs, palms, <i>Geinitzia</i>
	Phase 1 (Vermejo)	1	82	63	71	34	85	9	PDF	S Mg	Diverse lauraleans, diverse euphorbs, palms, <i>Geinitzia</i> , araucarian, 'Myrtophyllum'
Louisiana	Naborton	>5	47	44	83	75	91	63	TRF	W Mg	'Magnolia', lauraleans [needs restudy]

Numbers in parentheses are based on small samples and may not be reliable. Significant taxa are primarily those numerically common in the samples. Scientific names which are not correct, but for which no alternative has been proposed, are shown in quotation marks. Aff. indicates related taxon.

* A few previous collections were not examined; from taxa listed most physiognomic parameters can be calculated, but not leaf size index.

† NBLEF, notophyllous broad-leaved evergreen forest; PBLDF, polar broad-leaved deciduous forest; PDF, paratropical dry forest; PRF, paratropical rain forest; S, successional (probably mostly shrubby); SMDF, successional mesothermal deciduous forest; SPLDF, successional polar broad-leaved deciduous forest; SPRF, successional paratropical rain forest; TRF, tropical rain forest.

‡ D, dry; M, moist; W, wet; S, subhumid; Mg, megathermal; Ms, mesothermal.

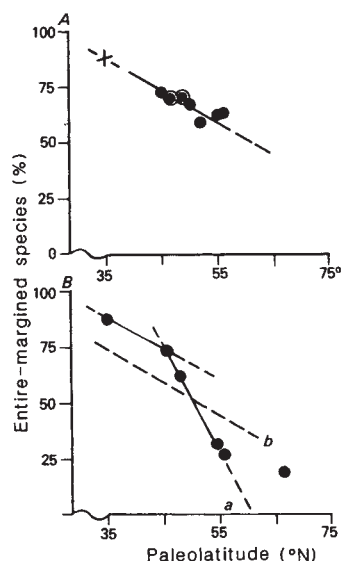


Fig. 3 Least-squares regression analyses of leaf-margin percentages for late Maestrichtian and B, early Palaeocene assemblages. In A, X denotes early Palaeocene Naborton assemblage and lies along the Maestrichtian regression line. In B, line *a* is based on the Raton through eastern Montana assemblages and line *b* is based on all assemblages. Only assemblages of >20 dicot species are included. A, $R^2 = 0.55$; B, *a*, $R^2 = 0.99$; *b*, $R^2 = 0.24$.

species may be analogous to an extant subgenus, genus or higher-level category²⁸. Survival of one species in a previously diverse genus would mask true species-level extinction.

The latitudinal pattern of extinction is that suggested for an 'impact winter'²⁴. Most (75%) Raton Lancian leaf species, especially Magnoliidae, are unknown in the early Palaeocene; dispersed cuticles of many of these taxa are found <1 cm below the K-T boundary¹⁶. Our analyses of the Denver floras indicate similarly high extinction. As at Raton, Denver deciduous taxa had far less extinction than evergreen taxa. In the northern high plains extinction across the K-T boundary was 50–60%²⁴. Notable in the early (and later) Palaeocene of the northern High Plains are deciduous taxa that were either uncommon in Lancian mesothermal vegetation or characteristic of polar deciduous forest (for example, trochodendroids and hamamelidaleans). Some evergreens apparently re-invaded from the south. Typically these taxa (for example, *Myrtophyllum*) are the few that occurred in both megathermal and mesothermal vegetation during the Lancian and thus may have survived in the south to re-expand northward temporarily; these taxa are typically rarer in the earliest Palaeocene than in the Lancian and unknown in the late early Palaeocene. Our preliminary analyses of known leaf floras in central Alberta indicate that most (76%) Lancian leaf lineages survived into the early Palaeocene. Extinction was highest among archaic deciduous and evergreen gymnosperms. However, the general pattern in the western interior is extinction of evergreen and survival of deciduous plants. Ability to enter dormancy was a major factor in survival across the K-T boundary.

In the Raton Basin vegetational change over an estimated 1.5 Myr mimics a successional pattern in diversity (quasisuccession)¹⁶. This pattern resulted from high extinction in megathermal evergreen vegetation at the K-T boundary. Mesothermal evergreen vegetation also suffered major extinction in the northern high plains, resulting in establishment and persistence of quasisuccessional, low-diversity deciduous forest analogous to successional vegetation in extant mesothermal regions. Polar deciduous forest was disrupted in the earliest Palaeocene but suffered little extinction and was re-established by the late early Palaeocene.

Occupation of the mesothermal region by deciduous forests

through most of the Palaeocene²⁹ indicates that there were no mesothermal refugia; otherwise, evergreens should have rapidly re-occupied this region. Replenishment of evergreens in mesothermal vegetation could then only occur by evolutionary adaptation from an impoverished megathermal flora. Slow replenishment of megathermal vegetation, evidenced by low diversity in Raton Phase 5 assemblages, suggests low survival rate in megathermal refugia.

High survival rate among deciduous versus evergreen plants indicates an event that favoured deciduous plants over evergreens sensitive to low temperature. The simplest explanation is a low-temperature excursion. High survival rate among aquatic ectotherms³⁰ places major constraints on severity and duration of the temperature excursion. Large aquatic ectotherms might survive a few months of a mean near freezing, but temperatures that would freeze water to a considerable depth, or a long period of inactivity without feeding would certainly be fatal (ref. 31 and J. H. Hutchison, personal communication). Many extant megathermal plants cannot tolerate temperatures <10 °C and most mesothermal broad-leaved evergreens cannot survive a cold-month mean <1 °C (ref. 10). Biological data are consistent with a short period (1–2 months) of a mean temperature near 0 °C. Deciduousness selected by the temperature excursion left a profound imprint on Northern Hemisphere Cenozoic vegetation^{29,32}. Rarity of deciduousness in Southern Hemisphere Cenozoic vegetation suggests that the temperature excursion had little effect in the Southern Hemisphere.

The sharpness and magnitude of vegetational and floral changes at the K-T boundary are dramatic. Mass-kill, ecological disruption and climatic change are associated with the event that produced the iridium-rich clay. The clays at various sites must be synchronous: were they of disparate ages, short-term successions, not quasisuccession, would have resulted. Although some³³ have proposed that the event was a marked and prolonged cooling from massive volcanic eruptions, the data from the land biota overall support a brief low-temperature excursion resulting from a bolide impact¹.

We thank W. A. Clemens, A. R. Sweet and officials of the Luscar-Sterco Mine for assistance in various aspects of fieldwork; T. M. Bown, W. A. Clemens, L. J. Hickey, A. H. Knoll, D. J. Nichols, C. B. Officer, P. H. Raven and R. A. Spicer for helpful discussions; and R. A. Stockey for access to the Genesee collections.

Received 11 April; accepted 2 September 1986.

1. Alvarez, W. *et al.* *Geol. Soc. Am. spec. Pap.* **190**, 305–315 (1982).
2. Alvarez, L. W. *et al.* *Science* **208**, 1095–1108 (1980).
3. Orth, C. J. *et al.* *Geol. Soc. Am. spec. Pap.* **190**, 423–433 (1982).
4. Pillmore, C. L. *et al.* *Science* **223**, 1180–1183 (1984).
5. Tschudy, R. H. *et al.* *Science* **225**, 1030–1032 (1984).
6. Bohor, B. F. *et al.* *Science* **224**, 867–869 (1984).
7. Lerbekmo, J. F. *et al.* *Can. J. Earth Sci.* (in the press).
8. Richards, P. W. *The Tropical Rain Forest* (Cambridge University Press, 1952).
9. Wolfe, J. A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **9**, 27–57 (1971).
10. Wolfe, J. A. *Prof. Pap. U.S. geol. Surv.* **1106**, (1979).
11. Wolfe, J. A. in *Paleobotany, Paleocology and Evolution* Vol. 2 (ed. Niklas, K. J.) 79–101 (Prager, New York, 1981).
12. Givnish, T. in *Topics in Plant Population Biology* (eds Solbrig, O. T., Jain, S., Johnson, G. B. & Raven, P. H.) 375–407 (Columbia University Press, New York, 1979).
13. Coley, P. D. *Ecol. Monogr.* **53**, 209–233 (1983).
14. Fahn, A. *Plant Anatomy* (Pergamon Press, Oxford, 1967).
15. Dorf, E. *Publ. Carnegie Instn* **508** (1942).
16. Wolfe, J. A. & Upchurch, G. R. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
17. Smith, A. G., Hurley, A. M. & Briden, J. C. *Phanerozoic Paleogeographic World Maps* (Cambridge University Press, 1981).
18. Wolfe, J. A. & Upchurch, G. R. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* (in the press).
19. Wolfe, J. A. *Am. geophys. Un. geophys. Monogr.* **32**, 357–375 (1985).
20. Batten, D. J., in *Fossils and Climate* (ed. Brechley, P.) 127–163 (Wiley, New York, 1984).
21. Upchurch, G. R. & Wolfe, J. A. in *Origins of Angiosperms and Their Biological Consequences* (eds Friis, E. M., Chaloner, G. W. & Crane, P. R.) (Cambridge University Press, in the press).
22. Bartels, W. S. *Pap. Univ. Mich. Mus. Paleont.* **24**, 73–79 (1980).
23. Hickey, L. J. *Pap. Univ. Mich. Mus. Paleont.* **24**, 33–49 (1980).
24. Hickey, L. J. *Nature* **292**, 529–531 (1981).
25. Wolfe, J. A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **30**, 313–323 (1980).
26. Erickson, J. M. A. *Proc. North Dakota Acad. Sci.* **32**, 79–89 (1978).
27. Boersma, A. in *Catastrophes and Earth History* (eds Berggren, W. A. & Van Couvering, J. A.) 247–277 (Princeton University Press, 1984).
28. Muller, J. *Biol. Rev. Cambridge phil. Soc.* **45**, 417–450 (1970).
29. Brown, R. W. *Prof. Pap. U.S. geol. Surv.* **375**, (1962).

30. Clemens, W. A. *Geol. Soc. Am. spec. Pap.* **190**, 407-413 (1982).
31. Hutchinson, J. H. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **37**, 149-164 (1982).
32. Wolfe, J. A. *Paleobiol.* (in the press).
33. Officer, C. B. & Drake, C. L. *Science* **227**, 1161-1167 (1985).
34. Lerbekmo, J. F., Evans, M. E. & Baadsgard, H. *Nature* **279**, 26-30 (1979).
35. Archibald, J. D. et al. *Geology* **10**, 153-159 (1982).
36. Nichols, D. J. & Ott, H. L. *Palynology* **2**, 94-112 (1978).
37. Hotton, C. *6th Int. Palynolog. Conf. Abstr.*, 66 (1984).
38. Sweet, A. R. & Hills, L. V. *6th Int. Palynol. Conf. Abstr.*, 160 (1984).
39. Middleton, M. D. *thesis*, Univ. Colorado (1983).
40. Knowlton, F. H. *Prof. Pap. US geol. Surv.* **130** (1922).
41. Knowlton, F. H. *Prof. Pap. US geol. Surv.* **155** (1930).
42. Berry, E. W. *Prof. Pap. US geol. Surv.* **185-F** (1934).
43. Brown, R. W. *Prof. Pap. US geol. Surv.* **189-I** (1939).
44. Bell, W. A. *Bull. geol. Surv. Can.* **13**, (1949).

Growth and longevity of krill during the first decade of pelagic whaling

A. A. Rosenberg, J. R. Beddington & M. Basson

IIED/IUCN Marine Resources Assessment Group, Imperial College, 48 Prince's Garden, London SW7 1LU, UK

Krill (*Euphausia superba* Dana), the dominant member of the zooplankton community of the Southern Ocean, has been the subject of study for many years. As the main prey species of the baleen whales, several species of seals and many sea bird, fish and cephalopod species, its importance within the ecosystem is clear. It is important for the practicalities of management of the Southern Ocean to estimate the ability of krill to sustain a harvest by man or by its natural predators^{1,2}. This estimation requires a knowledge of growth and longevity^{3,4}, which in krill are the subject of some controversy. Conflicting results have been obtained from krill size composition analysis, which indicates a short lifespan, and age pigment analysis, which indicates a longer life. Here we have analysed eight years of data on krill taken between 1928 and 1938 and have obtained estimates which consistently indicate high growth rates, modulated by a strong seasonal pattern. The cumulative growth through the year is therefore consistent with the studies of age structure using pigment analysis.

In the absence of any method for ageing krill early studies relied on the deduction of estimates of age from information obtained on the length composition of samples⁵⁻¹¹. The results of these studies indicated a high growth rate, with maximum length being approached within a short life cycle of 2-3 years. Recent work has used the age pigment lipofuscin to determine metabolic age^{12,13}. However, these studies suffer from a lack of knowledge of the relationship between metabolic age and real time. They have, to date, only been applied to very small samples of krill. The lipofuscin method, originally developed from studies on insects, has not yet addressed the difficulties of working with an animal like krill which does not have a terminal molt. An unknown proportion of the pigment is lost during ecdysis¹² which may affect the interpretation of results. Despite these problems, these studies indicate a rather longer life cycle with individuals of ages 6-7 being present. The implications for harvesting are that if the shorter lifespan is correct a much larger proportion of krill standing stock may be taken as harvest.

The most extensive time series of length composition samples of Antarctic krill were taken by the British R.R.S. Discovery expeditions which operated from 1926-1951. This sampling period coincided with the first period of major exploitation of the stocks of great whales in the Southern Ocean. We have re-analysed the length composition data on *Euphausia superba* for the period 1928-1938 with two recently developed, simple analytical methods to determine the growth of this species through the period and its expected longevity.

Length composition samples for the period 1928-1938 were taken from the logbooks of the Discovery expeditions and incorporated into a computer database by staff of the Institute of Oceanographic Sciences (Wormley, U.K.) and the Marine

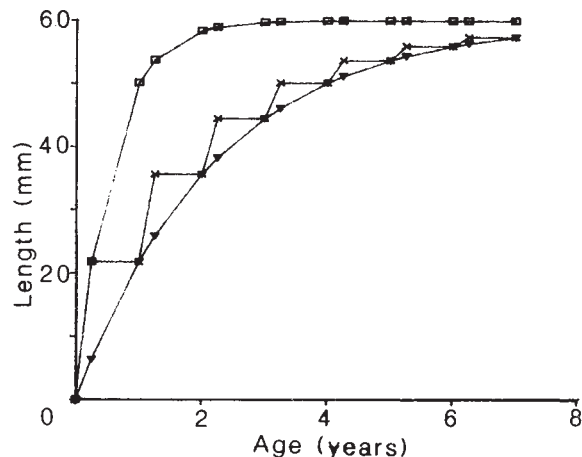


Fig. 1 Von Bertalanffy growth curves with parameters $\square$, $k = 1.8$ (equation 1); $\blacktriangledown$, $gk = 0.45$ (equation 2); and $\times$, the step function for seasonal growth with $k = 1.8$ and $g = 0.25$ (equation 3). In all cases $L_{inf} = 60.0$.

Resources Assessment Group of Imperial College (London, U.K.). The population structure of *E. superba* is unknown. Because of this we treated broad geographical regions as the units relevant for analysis. For each year the region containing the most extensive set of samples was used for the estimation of growth rate. We examined the data in two ways: as a standing stock, to provide an average view of the length composition over the year, and to find the monthly change in length composition through the growing season. With a few exceptions the samples were taken between September and May.

We consider growth to be described by the von Bertalanffy curve often used in fisheries biology. That is

$$l_t = L_{inf} (1 - e^{-kt}) \quad (1)$$

where l_t is the length of an animal at age t (where the age of the animal is measured from the time at which it had length zero), L_{inf} is a parameter giving the asymptotic maximum length for the animal and k is a growth constant with units $1/\text{time}$.

A dominant feature of the Southern Ocean environment is seasonality. There is substantial evidence that growth of many animals virtually ceases during the winter^{14,15}. Laboratory experiments on krill indicate that animals may even shrink if starved¹⁶. To account for such seasonality in samples which average the length composition over the year we modify equation (1) such that

$$l_t = L_{inf} (1 - e^{-gkt}) \quad (2)$$

where g is the proportion of the year in which the animals are growing. We hypothesize that the actual growth of the animal approximately follows a step function given by

$$l_t = L_{inf} (1 - e^{-k(t-n(1-g))}) \quad n < t \leq n+g$$

$$l_t = l_{n+g} \quad n+g < t < n+1 \quad (3)$$

where n is the number of winters the animal has survived. Figure 1 illustrates the growth curves given by equations (1) and (2) and the step function given by equation (3). The monthly length composition data were used to estimate the parameter k in equation (2). The data were the proportion of animals in the total monthly sample in each 5 mm length interval from 5 mm to 65 mm (see Fig. 2, top). To estimate k we write a projection matrix¹⁷, whose elements are a function of the von Bertalanffy equation parameters k and L_{inf} , which takes the length composition at time t to the composition at $t+dt$ where dt is the time between samples. This matrix is analogous to a Leslie population projection matrix. The first sample in a given growth season is projected forward to the sampling times of all subsequent length compositions taken in the same season using the matrix. The

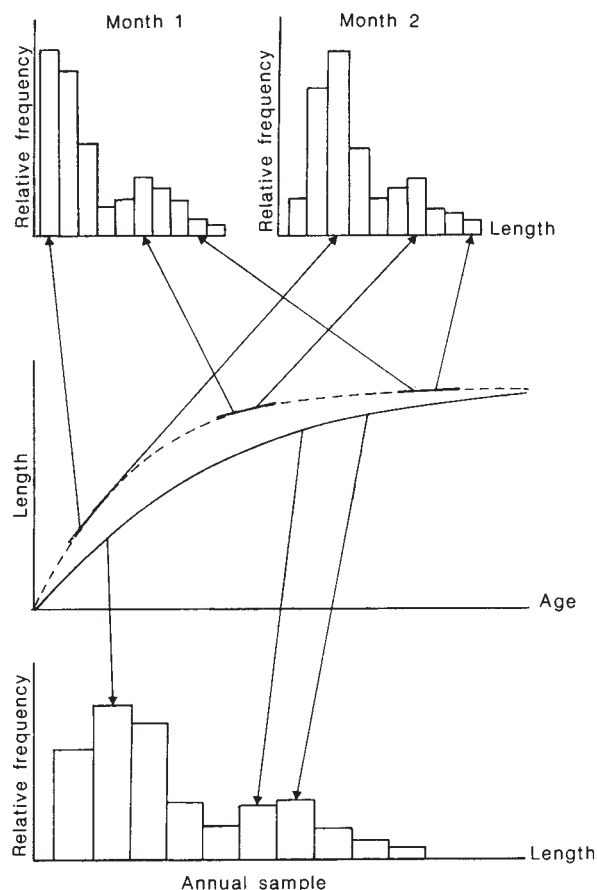


Fig. 2 Schematic representation of growth parameter estimation methods. Dashed line, growth curve for krill during growth season. Arrows indicate movement of individual krill from one population to another. Solid line, length against age for annual sample (age up to 6–7 years). The shape of the dashed line is estimated by examining the length frequency distributions in successive time periods during the growth season, (several months). The parameter k which best projects the timecourse of the samples through the season is chosen by least squares. The shape of the solid line is estimated using a method involving a scoring function that determines which parameter combination, gk , best match the modes observed in annual length compositions.

estimate of k is that value which best matches the predicted to observed length compositions under a least squares criterion (Fig. 2). The value of L_{inf} can be taken as the size of the largest individual observed assuming that the largest individuals in the population are sampled by the gear used. We set L_{inf} to 60 mm for the results reported here. However, our results for the estimation of k are insensitive to variations of 5 mm either side of this assumed value²⁰.

For the yearly (standing stock) length composition data the parameter estimates are approximately equivalent to the product gk in equation (2) since they represent growth averaged over the year, not just within the high growth season. We used a slight technical modification of Shepherd's^{18,20} method which scores the modes in the observed length composition based on their predicted position for a given growth curve (Fig. 2). Monte Carlo tests for the two methods indicate their accuracy to be within 10 and 15% of the true value of k respectively²⁰.

The estimated k values from the monthly data are quite consistent through the period (Table 1). The average is 1.80. This level of k gives a daily growth rate for krill during the fast growth season as 0.148 mm per day. Recent laboratory studies have shown *E. superba* to attain growth rates of around 0.13 mm per day^{15,19}.

The k values for the yearly data (Table 1) again show a high

Table 1 Growth parameter estimates for krill, *Euphausia superba*, sampled during the Discovery investigations

Split year (June–May)	Summer season growth		Yearly growth		Area	n
	k^*	mm per day†	$gk\ddagger$	g		
1928	2.18	0.179	0.43	0.20	Scotia Sea	751
1931	1.79	0.147	0.47	0.26	Scotia-Weddell	994
1934	1.42	0.117	0.27	0.19	Scotia-Weddell	594
1935	1.96	0.161	0.46	0.23	Australian	4,624
1936	1.28	0.105	0.47	0.37	Scotia-Weddell	10,208
1937	2.09	0.172	0.34	0.16	Australian	5,526
1938	1.89	0.155	0.47	0.25	Indian	6,148

Data are for split years over broad geographic regions in the Southern Ocean. Summer season growth refers to the von Bertalanffy parameter k estimated from a series of length frequency samples taken in a single growth season. Yearly growth is the estimated parameter combination, k times g , which represents the average growth over the entire year, not just the growing season. The number (n) of krill measured in each split year's samples which were used to estimate the length composition is given.

* Estimated from monthly length compositions—projection matrix.

† At a length of 30 mm growth rate in millimeters per day.

‡ Estimated from yearly length compositions—modal analysis.

degree of consistency, as with the monthly data. The estimated yearly k values are approximately 1/4 of those for the high growth season indicating that maximal growth occurs over approximately three months beginning in the early spring with the seasonal increase in primary production. Such strong seasonality in growth has been reported for other types of animals in Antarctic waters¹⁴. The strong seasonality in krill growth implies that modes in the length composition for each age group will be most distinct early in the austral spring and samples taken then be most valuable for monitoring growth rate.

Conflicting results concerning the possible lifespan of krill are now interpretable. Growth rates are quite high for this species as would be consistent with a short (2–3 yr) life cycle. However, since the animals only grow through a few months of the year, the effective rate over the entire year is more consistent with a 6–7 yr lifespan as estimated from the age pigment data. Put another way, the average effective growth rate estimated (represented by the step function in Fig. 1) means that it will take a minimum of 6–7 yrs for a krill to approach a length of 60 mm.

An important question is whether the growth of krill has changed in recent years in response to low levels of their baleen whale predators. Our results indicate growth to be quite variable between years and areas, but comparable analyses to our work on the Discovery data have not been performed on recent samples. The recently reported presence of 6–7 year old krill is not inconsistent with length compositions sampled during the first decade of pelagic whaling and so do not imply an increase in krill longevity in response to decreased predation.

This work was supported by the Natural Environment Research Council. A.A.R. was supported by a NATO Postdoctoral Fellowship. We thank Mr A. de C. Baker, Drs I. Everson & J. G. Shepherd for helpful advice and for help in setting up the database. The IIED/IUCN Marine Resources Assessment Group acknowledge support from IUCN and the World Wildlife Fund.

Received 24 April; accepted 10 September 1986.

1. Laws, R. M. *Phil. Trans. R. Soc. Lond. B* **279**, 81–96 (1977).
2. Beddington, J. R. & de la Mare, W. K. in *Marine Mammals and Fisheries* (eds Beddington, J. R., Beverton, R. J. H. & Lavigne, D. M.) 94–105 (George Allen & Unwin, London, 1985).
3. Gulland, J. A. *F.A.O. Fish. Biol. Tech. Pap.* **97**, 1–425 (1970).
4. Beddington, J. R. & Cooke, J. G. *F.A.O. Fish. Biol. Tech. Pap.* **242**, 1–47 (1983).
5. Ruud, J. T. *Halvård. Skr. Oslo* **2**, 1–105 (1932).
6. Bargmann, H. E. *Discovery Rep.* **23**, 103–176 (1945).
7. Marr, J. *Discovery Rep.* **32**, 33–464 (1962).
8. Makarov, R. R., Naumov, A. G. & Shevtsov, V. V. in *Antarctic Ecology* Vol. 1 (ed. Holdgate, M. W.) 173–176 (Academic, London, 1970).
9. Ivanov, B. G. *Marine Biol.* **7**, 340–351 (1970).
10. MacKintosh, N. A. *Discovery Rep.* **36**, 1–94 (1972).
11. MacKintosh, N. A. *Discovery Rep.* **36**, 95–156 (1973).
12. Ettershank, G. *Polar Biol.* **2**, 189–193 (1983).
13. Ettershank, G. *J. Crustacean Biol.* **4** (Spec. No. 1), 295–305 (1984).

14. Clark, A. *Mar. Biol. A. Rev.* **21**, 341-453 (1983).
15. Bucholtz, F. in *Antarctic nutrient cycles and food webs* (eds Siegfried, W. R., Condy, P. R. & Laws, R. M.) 339-345 (Springer, Berlin, 1985).
16. Ikeda, T. & Dixon, P. J. *exp. mar. Biol. Ecol.* **62**, 143-151 (1982).
17. Shepherd, J. G. in *The theory and application of length based methods of stock assessment* (eds Pauly, D. & Morgan, G. P.) Section 1 (ICLARM Conference Series, Manila, in the press).
18. Shepherd, J. G. in *The theory and application of length based methods of stock assessment* (eds Pauly, D. & Morgan, G. P.) Section 2 (ICLARM Conference Series, Manila, in the press).
19. Morris, D. J. & A. Keck *Meeresforschung* **30**, 94-100 (1984).
20. Basson, M., Rosenberg, A. A. & Beddington, J. R. (manuscript in preparation).

Ocular dominance shift in kitten visual cortex caused by imbalance in retinal electrical activity

Barbara Chapman, Michael D. Jacobson,
Holger O. Reiter & Michael P. Stryker

Division of Neurosciences and Department of Physiology,
University of California, San Francisco, California 94143, USA

Monocular lid suture during the sensitive period early in the life of a kitten disrupts normal development of inputs from the two eyes to the visual cortex, causing a decrease in the fraction of cortical cells responding to the deprived eye¹. Such an ocular dominance shift has been assumed to depend on patterned visual experience, because no change in cortical physiology is produced by inequalities between the two eyes in retinal illumination² or temporally modulated diffuse light stimulation^{3,4}. A higher-level process, involving gating signals from areas outside striate cortex, has been proposed to ensure that sustained changes in synaptic efficacy occur only in response to behaviourally significant visual inputs⁵. To test whether such a process is necessary for ocular dominance plasticity, we treated 4-week-old kittens with visual deprivation and monocular tetrodotoxin (TTX) injections to create an imbalance in the electrical activities of the two retinas in the absence of patterned vision. After 1 week of treatment we determined the ocular dominance distribution of single units in primary visual cortex. In all kittens studied, a significant ocular dominance shift was found. In addition to this physiological change, there was an anatomical change in the lateral geniculate nucleus, where cells were larger in laminae receiving input from the more active eye. Our results indicate that patterned vision is not necessary for visual cortical plasticity, and that an imbalance in spontaneous retinal activity alone can produce a significant ocular dominance shift.

For this study 11 28-34-day-old kittens received intravitreal injections of 5 mM TTX in one eye to block ganglion-cell activity in that eye⁶. Patterned visual input to the other eye was eliminated by suturing the eyelid. Six of the kittens (TTX/lid-suture group) were kept in a normal light cycle, and thus experienced no activity in one eye and diffuse, temporally modulated light in the other. The remaining five kittens (TTX/dark group) were kept in the dark, experiencing no activity in one eye and only spontaneous activity in the other eye. The retinal action potential blockade was maintained for 1 week by repeating the TTX injections. The lid of the injected eye was then sutured and the animals remained or were placed in the dark to prevent vision during recovery of retinal activity. Twenty-four to 48 hours after the retinal blockade began to wear off, the kittens were prepared for electrophysiological recording as described elsewhere⁷.

To demonstrate that normal activity had returned to the previously blocked eye, 5-13 single units with central receptive fields (within 15° of the area centralis) were recorded in the lateral geniculate nucleus (LGN) contralateral to the injected eye. In addition, multi-unit responses were assessed every 100 µm through laminae A and A1. No difference in responsiveness was detected between cells driven by the two eyes (see Figs



Fig. 1 Coronal Nissl-stained section through an LGN contralateral to the TTX-injected eye. The section shown is 2.88 mm caudal to the rostral pole of the LGN. An electrode track is shown with electrolytic marking lesions indicated by curved arrows. Filled arrows, locations of single units responding to the untreated ipsilateral eye; open arrows, single units responding to the previously blocked contralateral eye. Receptive fields of the units indicated by the top two arrows were at elevation 30°, azimuth 9°; those of the units indicated by the lower five arrows ranged between elevation 6°, azimuth 4° and elevation 4°, azimuth 2°. Scale bar, 0.5 mm.

1, 2), nor between multi-unit recordings in laminae responsive to the two eyes. Briskly responsive ON- and OFF-centre and sustained and transient units were recorded in both laminae.

To determine whether an ocular dominance shift had occurred, we studied cortical area 17 contralateral to the injected eye. This hemisphere was chosen because normal kittens show an ocular dominance bias toward the contralateral eye⁶; thus if recording was in the hemisphere ipsilateral to the TTX-treated eye, an apparent shift in ocular dominance towards the more active eye could be caused, at least in part, by this bias. In the hemisphere contralateral to the injected eye, however, an ocular dominance distribution favouring the more active (ipsilateral) eye must be caused by the effects of deprivation. Horizontal electrode penetrations were used to sample cells across, rather than down, ocular dominance columns; each penetration was long enough to pass through several columns. The intended sample size was 30 cells per animal. Receptive fields of all units studied were within 15° of the area centralis. Ocular dominance was determined for each visually responsive unit using the standard seven-point scale⁸. For both conditions, TTX/lid suture and TTX/dark, ocular dominance histograms were significantly shifted toward the more active, ipsilateral eye (Fig. 3b, c). These data contrast with histograms from normal kittens of the same age, which show a slight bias in favour of the contralateral eye (Fig. 3a), as well as with the more strongly shifted histogram from conventionally monocularly lid-sutured kittens (Fig. 3d). The degree of shift in each kitten, estimated by a scalar index, indicates that our two groups are significantly different from both normal and conventionally monocularly deprived animals (Mann-Whitney *U* test, $P < 0.01$), but that the results in the TTX/lid-suture and TTX/dark groups are not significantly different (Mann-Whitney *U* test, $P > 0.05$) (Fig. 4a).

In 7 of the original 11 kittens, cross-sectional areas of cells in the LGN were measured in Nissl-stained sections. (LGN cell sizes are known to be smaller in the laminae responding to the deprived eye in kittens with monocular lid suture^{1,9} or monocular TTX blockade¹⁰.) As expected, cell sizes were smaller in laminae responding to the TTX-treated eye in both TTX/lid-suture and TTX/dark animals. These morphological results, like the physiological ones, are intermediate between those from normal and monocularly deprived animals (Fig. 4b).

Although unlikely, the changes in cortical physiology observed could have been caused by damage to the TTX-treated eye that was not detected in our geniculate recordings, rather

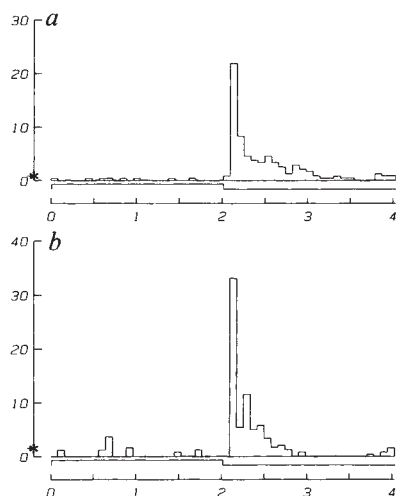


Fig. 2 Post-stimulus histograms showing examples of the flash response of single OFF-centre units driven by the untreated ipsilateral eye (*a*) and by the previously blocked contralateral eye (*b*) in the same LGN. Stimulus intensity was one log unit above threshold. *y*-axis, spikes per s; *x*-axis s. Asterisk, level of spontaneous activity; stimulus time course, line above the *x*-axis (ON, up; OFF, down).

than by the activity blockade. If such toxicity was responsible for the ocular dominance shifts, the effects should be permanent. If, however, differential activity were responsible, then the shifts should be reversible, as are those produced by conventional monocular deprivation in kittens of the same age¹¹. To test the reversibility of our treatment, two additional kittens were studied. These kittens experienced 1 week of the TTX/lid-suture treatment starting at 28 or 32 days old. After the TTX blockade wore off, the kittens remained in a normal light cycle for 9 or 11 days with the sutured eyelid remaining closed. Thus, these kittens experienced a week of no activity in the TTX-treated eye combined with diffuse light vision through a closed lid in the other eye, followed by a period of conventional monocular deprivation with patterned vision in the previously TTX-treated eye. The first week of treatment presumably produced an ocular dominance shift toward the lid-sutured eye, as occurred in TTX/lid-suture animals. If TTX injections did not cause permanent damage, this shift should have been reversed during the period of monocular deprivation, resulting in a cortical bias in favour of the eye that had been blocked during the first part of the treatment and had experienced patterned vision during the second part. Single-unit recordings from area 17 of the two kittens showed that such a reversal had occurred. Ocular dominance histograms from both kittens were strongly shifted toward the eye that was first blocked and then had patterned vision. The degree of shift in these kittens (contralateral bias indices 86 and 99; see Fig. 4 legend) was comparable to that seen in conventional reverse lid-sutured kittens of the same age¹¹. Thus TTX injections do not cause lasting damage to the ability of the treated eye to compete for cortical dominance.

The similarity of three of the present findings to those from conventional monocular deprivation experiments indicates that the ocular dominance shifts seen in TTX/lid-suture and TTX/dark animals were caused by the differences between activities of the two eyes, rather than by damage to the retinal ganglion cells of the injected eye. (1) LGN responses to stimulation of the two eyes were equally good even in the most shifted animals in which few or no cortical cells responded to the injected eye. (2) LGN cell sizes were smaller in laminae responding to the injected eye. (3) The ocular dominance shifts produced were readily reversible.

We draw three conclusions from the results of these experiments. First, the significant changes in ocular dominance in both TTX/lid-suture and TTX/dark animals show that patterned

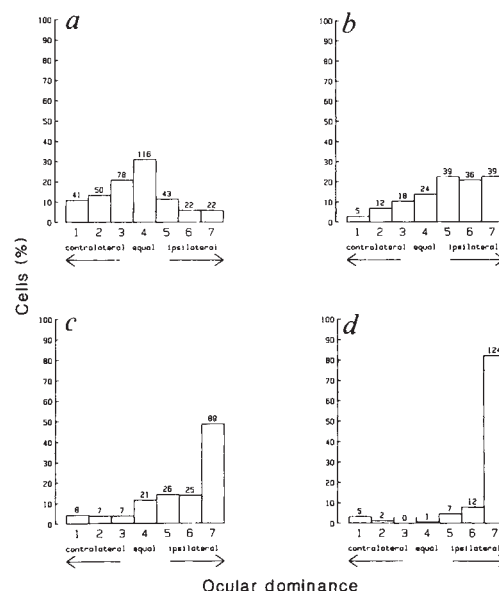


Fig. 3 Ocular dominance histograms⁸ compiled from single unit responses in area 17. An ocular dominance of 1 indicates a cell driven only by the contralateral eye; 7 a cell driven only by the ipsilateral eye; and 4 a cell driven equally by the two eyes. In each deprivation condition, the contralateral eye is less active. *a*, 372 visually responsive units from six normal kittens, 36–51 days old: 26 visually unresponsive units recorded⁶. *b*, 173 visually responsive units from five TTX/dark kittens, 37–44 days old: 12 visually unresponsive units recorded. *c*, 183 visually responsive units from six TTX/lid-suture kittens, 36–40 days old: 19 visually unresponsive units recorded. *d*, 151 units from four kittens monocularly deprived by lid suture for 3.5–6 days, followed by 0–2 days recovery in darkness; 29–34 days old^{9,15,16}.

vision is not necessary to produce changes in cortical physiology. Second, the results from the TTX/dark group indicate that an imbalance in spontaneous activity alone is enough to cause an ocular dominance shift. Third, the similar results produced by TTX/dark and TTX/lid-suture treatments suggest that the vision that occurs through a closed eyelid has little additional effect on ocular dominance.

It is generally believed that cortical plasticity is dependent on cortical cells being driven by their geniculate inputs. Silencing all cortical activity through intra-cortical TTX injections prevents ocular dominance shifts in monocularly deprived animals¹². Because geniculate¹³ and cortical¹⁴ spontaneous activities are dependent on levels of spontaneous activity in the retina, spontaneous activity must drive cortical cells, however weakly. Thus the present results are consistent with a model of plasticity requiring that cortical cells be driven.

Our conclusion that ocular dominance shifts can be produced in the absence of patterned vision does not necessarily refute the central gating hypothesis⁵. Although shifts in our animals did occur without behaviourally significant visual input to either eye, these shifts were not as large as those seen in conventionally monocularly deprived animals^{9,15,16}, although the differences in retinal activities were extreme. Perhaps the central gate is always slightly open, allowing some degree of shift to occur. Behaviourally significant information could further open the gate, causing the larger ocular dominance shifts seen in conventionally monocularly deprived animals. We do not know why ocular dominance shifts were produced by the present conditions but not by other rearing conditions with different inputs to the two eyes in the absence of patterned vision^{2–4}. The actual retinal activities produced by any of these rearing conditions are, however, unknown (except that TTX does block all activity⁶). Therefore, either a large quantitative difference or some temporally modulated qualitative difference between the

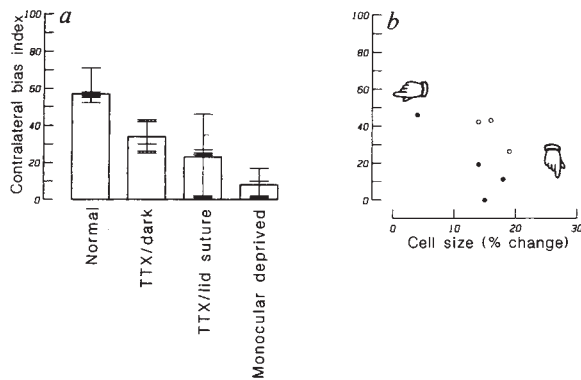


Fig. 4 *a*, Contralateral bias indices showing the range of findings among individual animals from Fig 3. *b*, Scatter plot of per cent change in geniculate cell size versus contralateral bias index. *a*, The index provides a single number reflecting the degree of ocular dominance shift for each animal. The index is calculated according to the equation

$$100[(1-7)+2/3(2-6)+1/3(3-5)+n]/2n$$

Italicized numbers, number of units in each ocular dominance group; n , total number of visually responsive units. This index is 100 if all cells are driven exclusively by the contralateral eye and 0 if all cells are driven exclusively by the ipsilateral eye. The greater variability among TTX/lid-suture animals is unexplained; it may be caused by differences among the animals of this group in the amount of diffuse-light stimulation of the retina of their sutured eyes. *b*, Cells (75 per lamina) were measured in both LGNs of each animal. Measurements were made in the latero-medial middle third of coronal sections taken ~2.5 mm caudal to the rostral pole of the nucleus (the area corresponding to the central 5–10° of the visual field)²⁶. All cell bodies with visible nucleoli were traced in the camera lucida under $\times 63$ or $\times 100$ oil-immersion objectives at final magnifications of $\times 1,160$ – $1,840$. The areas of these tracings were then measured using a Summagraphics Bitpad Planimeter. The standard errors ranged from 4 to 6% of the mean for individual laminae. Per cent change was calculated as $1 - \frac{\sum \text{mean cell sizes from laminae responding to the injected eye}}{\sum \text{mean cell sizes from laminae responding to the more active eye}}$. Open circles, data from TTX/lid-suture animals; filled circles, data from TTX/dark animals. Upper finger, region of data from normal animals⁶; lower finger, region of data from monocularly deprived animals⁹. The spread of these control data along the ordinate is shown in *a*; the spread along the abscissa is ~4% for normal animals and 7% for monocularly deprived animals⁹. Although there is no strong correlation between the change in cell size and the degree of ocular dominance shift in our animals, the extents of both morphological and physiological changes are intermediate between data from normal and conventional monocularly deprived animals. The morphological results for TTX/dark animals (average percentage change = 16) are very similar to those previously reported for older TTX/dark kittens (average percentage change = 17)¹⁰.

activities of the two retinas may be necessary to produce ocular dominance shifts.

Our present finding that imbalances in spontaneous activity alone can cause changes in connections to cortical cells, together with earlier work on the effects of spontaneous activity blockade on the development of ocular dominance columns^{6,17}, is consistent with the hypothesis that spontaneous electrical activity plays a part in the normal fetal development of the visual system¹⁸. Activity-dependent rearrangements of ocular dominance columns, which occur postnatally in the cat¹⁹, begin 3–6 weeks prenatally in the monkey^{20,21}. If the development of ocular dominance columns follows similar principles in the two species²², it is not surprising that spontaneous activity plays this part in the cat because in the monkey *in utero*, vision is occluded, and the only possible electrical activity is spontaneous. Observations that retinogeniculate axons make functionally effective binocular connections prenatally²³, some of which are lost before birth^{23,24}, suggest that spontaneous activity is also involved in the refinement of these connections^{24,25}.

Received 28 January; accepted 23 July 1986.

- Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **26**, 1003–1017 (1963).
- Blakemore, C. *J. Physiol., Lond.* **261**, 423–444 (1976).
- Singer, W., Rauschecker, J. & Werth, R. *Brain Res.* **134**, 568–572 (1977).
- Wilson, J. R., Webb, S. V. & Sherman, S. M. *Brain Res.* **136**, 277–287 (1977).
- Singer, W. in *The Neurosciences Fourth Study Program* (eds Schmitt, F. O. & Worden, F. G.) 1093–1110 (MIT, Cambridge, 1979).
- Stryker, M. P. & Harris, W. J. *Neurosci.* **6**, 2117–2133 (1986).
- Schopmann, A. & Stryker, M. P. *Nature* **293**, 574–576 (1981).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160**, 106–154 (1962).
- Movshon, J. A. & Dursteler, M. R. *J. Neurophysiol.* **40**, 1255–1265 (1977).
- Kupfermann, B. D. & Kasamatsu, T. *Nature* **306**, 465–468 (1983).
- Movshon, J. A. *J. Physiol., Lond.* **261**, 125–174 (1976).
- Reiter, H. O., Waitzman, D. M. & Stryker, M. P. *Soc. Neurosci. Abstr.* **11**, 463 (1985).
- Eysel, U. T. & Wolfherd, E. *J. comp. Neurol.* **229**, 301–309 (1983).
- Harris, W. A. & Stryker, M. P. *Soc. Neurosci. Abstr.* **3**, 1785 (1977).
- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **206**, 419–436 (1970).
- Olson, C. R. & Freeman, R. D. *J. Neurophysiol.* **38**, 26–32 (1975).
- Stryker, M. P. *Soc. Neurosci. Abstr.* **7**, 842 (1981).
- Stryker, M. P. in *Developmental Neurophysiology* (eds Kellaway, P. & Purpura, D.) (Johns Hopkins University, in the press).
- LeVay, S. & Stryker, M. P. in *Aspects of Developmental Neurobiology* (ed. Ferrendelli, J. A.) 83–96 (Soc. Neurosci., 1979).
- LeVay, S., Wiesel, T. N. & Hubel, D. H. *J. comp. Neurol.* **191**, 1–51 (1980).
- Rakic, P. *Nature* **261**, 467–471 (1976).
- Des Rosiers, M. H. *et al. Science* **200**, 447–449 (1978).
- Shatz, C. & Kirkwood, P. *J. Neurosci.* **4**, 1378–1397 (1984).
- Hubel, M. W., Stark, L. A. & Archer, S. M. *J. Neurosci.* **6**, 1021–1036 (1986).
- Sur, M., Garraghty, P. E. & Stryker, M. P. *Soc. Neurosci. Abstr.* **11**, 805 (1985).
- Sanderson, K. J. *J. comp. Neurol.* **143**, 101–118 (1971).

Intracellular calcium ions decrease the affinity of the GABA receptor

Inoue, M., Oomura, Y., Yakushiji, T. & Akaike, N.*

Department of Physiology, Faculty of Medicine, Kyushu University 60, Fukuoka 812, Japan

Intracellular free Ca^{2+} ($[\text{Ca}^{2+}]_i$) plays a crucial role in the transduction of extracellular signals¹. It has been implicated in the modulation of light sensitivity in *Limulus* photoreceptors² and in the efficacy of synaptic transmission^{3–5}; calcium ion fluxes are also involved in the postsynaptic facilitation of nicotinic transmission seen in sympathetic ganglia, and in activation of the acetylcholine (ACh) receptor^{6,7}. $[\text{Ca}^{2+}]_i$ is also a second messenger for many biologically active substances⁸. We recorded neuronal activities of sensory neurones from the bullfrog (*Rana catesbiana*), using the suction pipette method⁹ and a 'concentration clamp' technique¹⁰ to apply γ -aminobutyric acid (GABA) to the cell. We report the first evidence that $[\text{Ca}^{2+}]_i$ suppresses the GABA-activated Cl^- conductance, by decreasing the apparent affinity of the GABA receptor.

Sensory neurones were isolated from dorsal root ganglia by both mechanical dissection and enzymatic digestion in normal Ringer solution containing 0.3% collagenase and 0.05% trypsin at 37 °C for 15 min. The isolated cells were maintained in a solution containing equal parts of Ringer and Eagle's minimum essential medium at room temperature (~22 °C). The concentration clamp technique enables complete exchange of the external solution around the cell to occur within a few ms (ref. 10). K^+ and Na^+ in both internal and external solutions were replaced with equimolar Cs^+ and Tris, respectively, unless otherwise stated. All experiments were performed at room temperature. The data have been analysed statistically using Student's *t*-test.

Under current clamp conditions, rapid application of GABA (10^{-5} M) in the absence of external Ca^{2+} evoked a large depolarization, and a relatively stable plateau was reached within 2 sec (Fig. 1a). This response is mediated through an increase in Cl^- conductance¹¹. When the application of GABA was preceded by a series of Na^+ spikes (Fig. 1b), no effect on the GABA response was apparent. However, elicitation of a single Ca^{2+} spike in medium containing Ca^{2+} markedly reduced

* To whom correspondence should be addressed.

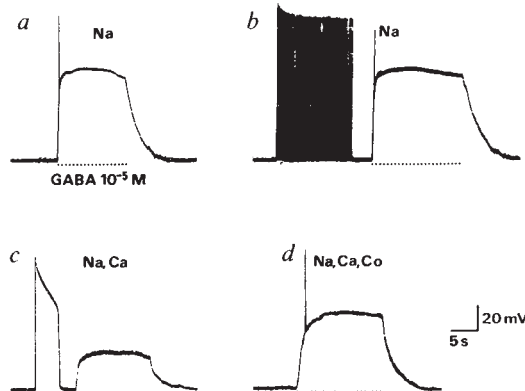


Fig. 1 Effects of Ca^{2+} spike and Na^{+} spikes on the GABA-induced depolarization. *a*, 10^{-5} M GABA depolarized the membrane by 75 mV in Na^{+} solution (in mM): NaCl 83, MgCl_2 5, TEA-Cl 25, CsCl 2 and glucose 5. The membrane potential was held at -80 mV. *b*, The preceding 50 Na^{+} spikes elicited by depolarizing pulses at a frequency of 3.3 Hz had little influence on the 10^{-5} M GABA response (74 mV) in an external Na^{+} solution. *c*, Marked suppression of 10^{-5} M GABA response after a 4.3 s duration Ca^{2+} spike in a Na^{+} and Ca^{2+} external solution where 2 mM CaCl_2 replaced 5 mM MgCl_2 and 6 mM NaCl was added to equalize osmolarities; depolarization of 32 mV. *d*, With suppression of the Ca^{2+} spike by adding 4 mM CoCl_2 to the Na^{+} and Ca^{2+} external solution, 10^{-5} M GABA produced a depolarization of 62 mV which was slightly smaller than the control, probably due to Co^{2+} interference. GABA equilibrium potential of 4 mV agreed well with the theoretical value estimated from the ionic activities of Cl^{-} in the external and internal solutions. This result suggests that the internal perfusion was complete. The ionic compositions of the standard solutions were (mM): internal, CsCl 95, (TEA-Cl) 25, Cs-aspartate 10, 0.5 EGTA- Ca buffer ($[\text{Ca}^{2+}]_i = 3 \times 10^{-8}$ M) and pH 7.2; external, Tris-Cl 89, TEA-Cl 25, CsCl 2, CaCl_2 2, glucose 5 and pH 7.4. The pH of all solutions was adjusted with Tris and HEPES buffers.

the subsequent response to GABA (Fig. 1c). This inhibitory effect was not simply due to the presence of Ca^{2+} in the external perfusate, since it could be reversed by adding Co^{2+} , which blocks the voltage-dependent Ca^{2+} current (I_{Ca})¹² (Fig. 1d).

Reduction of the GABA response by the Ca^{2+} influx was quantified in Na^{+} , K^{+} -free external solution containing 2 mM Ca^{2+} . Cells were voltage clamped to a holding potential of -50 mV and I_{Ca} was evoked by 50 mV depolarizing command pulses of varying duration. As can be seen in Fig. 2a, the GABA responses were suppressed to a greater extent when the Ca^{2+} influx was increased. A typical relationship between the total amount of I_{Ca} ($\int I_{\text{Ca}}$) and the GABA-induced peak response is

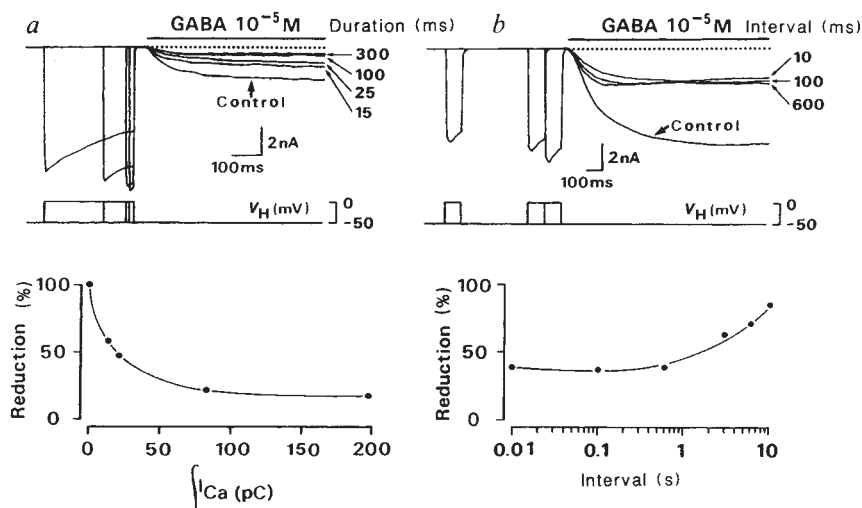
illustrated in the lower graph of Fig. 2a. It is evident that I_{Cl} in this cell is maximally inhibited when the Ca^{2+} influx exceeds 80 pC. The degree of this I_{Ca} -mediated suppression of I_{Cl} depended on the chelating potency of the internal solution used. When cells were perfused with a solution containing 2.5 mM EGTA, the subsequent response to 10^{-5} M GABA was reduced to 75% ($n = 15$) of control by the preceding 100 ms I_{Ca} , whereas in neurones perfused with 0.5 mM EGTA- Ca^{2+} buffer ($[\text{Ca}^{2+}]_i \approx 3 \times 10^{-8}$ M), the I_{Cl} was reduced to 50% ($n = 14$) ($p < 0.01$).

An increase of $[\text{Ca}^{2+}]_i$ facilitates the onset of depolarization in response to ACh in the frog endplate¹³. In the present study, a similar phenomenon was observed when the time interval between I_{Ca} and the GABA response was short (10 ms as shown in Fig. 2b). In this case, the GABA response reached a maximal value within 250 ms and thereafter I_{Cl} decreased gradually, whereas in the control it continued to rise. However, this acceleration of the GABA-response decay by I_{Ca} was inversely related to the time interval between I_{Ca} and GABA application and was not detectable with an interval of 600 ms, indicating that this facilitated desensitization after an I_{Ca} is only 'apparent'. This phenomenon could be accounted for by the time-dependent activation of the suppressive mechanism triggered by an increase in $[\text{Ca}^{2+}]_i$. Figure 2 shows that following I_{Ca} , full inhibition was only achieved after a few hundred milliseconds. The lower graph in Fig. 2b illustrates that recovery from the Ca^{2+} -mediated inhibition occurs within a few tens of seconds. It is not certain whether this time course is one of a change in $[\text{Ca}^{2+}]_i$ after I_{Ca} or release of Ca^{2+} from the receptive sites of GABA receptor complexes or a combination of both. In experiments with Ca^{2+} indicator dyes, increases in $[\text{Ca}^{2+}]_i$ are reported to last for over 10 s after 1 s I_{Ca} under voltage clamp conditions in sympathetic neurones¹⁴. These data, however, should not be directly compared with our results, because we recorded neuronal activity with a suction electrode.

It is probable that by dialysis, we removed the soluble cytoplasmic constituents that are important for physiological Ca^{2+} sequestration. However, this disadvantage is not expected to be significant in Ca^{2+} buffering, because in our experiments, cells were perfused with EGTA-containing solutions.

It has been reported that ACh receptor/ Cl^{-} channel complexes in *Aplysia* neurones are inactivated by intracellular Ca^{2+} binding to the receptor complexes¹⁵. However, Ca^{2+} -mediated inhibition of the Cl^{-} response in this preparation did not depend on the concentration of ACh, a result which markedly differs from ours. As shown in Fig. 3, the GABA-activated I_{Cl} was inhibited in a competitive manner by I_{Ca} . The degree of this suppression did not depend on membrane potential (which ranged from -40 to -100 mV), although the Ca^{2+} -dependent Cl^{-} or K^{+} conductance^{16,17} is markedly voltage-dependent.

Fig. 2 Suppressing action of I_{Ca} on the GABA-activated I_{Cl} . *a*, Top, actual records of 10^{-5} M GABA-gated I_{Cl} , with and without a preceding I_{Ca} elicited by the depolarizing pulse from the holding potential (V_H) of -50 mV to 0 mV for various durations (15, 25, 100 and 300 ms). The leakage and capacitive currents associated with the ionic currents were subtracted during the experiments by adding the current responses to equal but opposite voltage steps using a signal averager. Bottom, relationship between GABA-gated I_{Cl} and level of voltage-dependent Ca^{2+} influx (pC). The suppression of GABA-gated I_{Cl} reached a plateau with increasing amounts of I_{Ca} . *b*, Top, actual records of 10^{-5} M GABA-gated I_{Cl} elicited at various intervals after I_{Ca} . Note that the rising phase of I_{Cl} was slow and 'desensitization' became obscure with increasingly large intervals. Bottom, suppression of GABA responses plotted against the interval between the I_{Ca} and GABA responses.



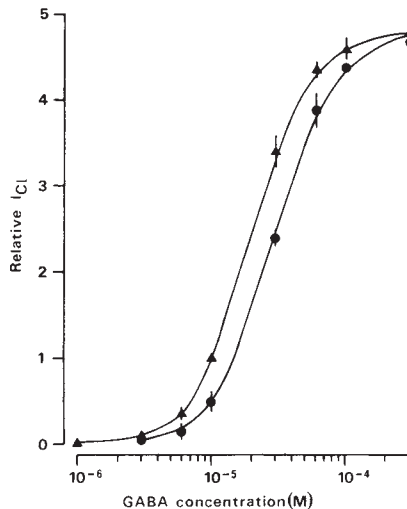


Fig. 3 Effect of 100 ms I_{Ca} on the GABA dose-response curve. The control dose-response curve (Δ) was made by averaging 8 cases, yielding a good fit to the theoretical expression assuming a Hill number of 2 and K_a of 2×10^{-5} M, corresponding to the concentration of GABA producing a half-maximum response. When each GABA response was preceded by a 100 ms I_{Ca} ($\bullet$), more than 5 cases at each concentration were averaged and the data points fell on the line assuming a Hill number of 2 and K_a of 3×10^{-5} M. Each data point represents mean $\pm$ s.e.

These results suggest that suppression of the GABA response is related to a decrease in the affinity of the GABA receptor for its agonist. We cannot exclude, however, other possibilities such as changes in channel gating kinetics.

The experiments reported here clearly demonstrate that a preceding I_{Ca} markedly suppresses the GABA-activated I_{Ci} , and the inhibitory effects are influenced by the EGTA concentration in the internal solution. Therefore, it is reasonable to assume that this suppression may be due to an increase in $[Ca^{2+}]_i$ subsequent to the influx of Ca^{2+} . However, I_{Ca} could still inhibit the GABA responses in cells perfused with an internal solution containing up to 2.5 mM EGTA. These observations suggest that inhibition of the GABA-activated I_{Ci} by Ca^{2+} might occur at a site that EGTA can not gain access to, as in Ca^{2+} -dependent K^+ conductance in smooth muscle¹⁸.

Biochemical studies^{17,18} have revealed that various endogenous substances that reduce the binding of GABA are present in the synaptic membranes of the mammalian central nervous system. The relationship between Ca^{2+} -mediated inhibition and these substances is currently under investigation. The suppression of GABA sensitivity by $[Ca^{2+}]_i$ may be of great importance in phenomena such as epileptogenesis, where the significant Ca^{2+} entry associated with paroxysmal depolarization shifts, may reduce the effectiveness of recurrent inhibition and therefore both prolong burst discharge and enhance the synchronization of firing. To our knowledge, this is the first demonstration that $[Ca^{2+}]_i$ rapidly (within a hundred ms) controls the apparent affinity of the GABA receptor.

This study was supported in part by grants 59123005 and 59480107 from The Ministry of Education, Science and Culture (N.A.). We thank M. Ohara (Kyushu University) and J. Behrends for comments on the manuscript.

9. Akaike, N., Hattori, K., Inomata, N. & Oomura, Y. *J. Physiol., Lond.* **360**, 367-386 (1985).
10. Akaike, N., Inoue, M. & Krishtal, O. A. *J. Physiol., Lond.* (in the press).
11. Hattori, K., Akaike, N., Oomura, Y. & Kuraoka, S. *Am. J. Physiol.* **246**, C259-C265 (1984).
12. Ishizuka, S., Hattori, K. & Akaike, N. *J. Membrane Biol.* **78**, 19-28 (1984).
13. Miledi, R. *Proc. R. Soc. Lond. B209*, 447-452 (1980).
14. Smith, S. J., MacDermott, A. B. & Weight, F. F. *Nature* **304**, 350-352 (1983).
15. Chemeris, N. K., Kazachenko, V. N., Kislov, A. N. & Kurchikov, A. L. *J. Physiol. Lond.* **323**, 1-19 (1982).
16. Miledi, R. & Parker, I. *J. Physiol. Lond.* **357**, 173-183 (1984).
17. Meech, R. W. & Standen, N. B. *J. Physiol. Lond.* **249**, 211-239 (1975).
18. Mitra, R. & Morad, M. *Science* **229**, 269-272 (1985).
19. Guidotti, A. et al. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6074-6088 (1982).
20. Yoneda, Y. & Kuriyama, K. *Nature* **285**, 670-673 (1980).

The t(8; 14) chromosomal translocation occurring in B-cell malignancies results from mistakes in V-D-J joining

Frank G. Haluska, Sheldon Finver, Yoshihide Tsujimoto & Carlo M. Croce

The Wistar Institute, 3601 Spruce Street, Philadelphia, Pennsylvania 19104, USA

The reciprocal chromosome translocation, t(8; 14), involving the heavy chain locus on chromosome 14 (refs 1, 2) and the *c-myc* oncogene on chromosome 8 (refs 3-5) is a characteristic of the B-cell malignancies Burkitt's lymphoma⁶ and acute lymphoblastic leukaemia (ALL)⁷. We have cloned and sequenced the t(8; 14) breakpoints of an African Burkitt's lymphoma cell line, P3HR-1, and a pre-B cell ALL cell line, 380 (ref. 8). In each case the region of chromosome 8 involved has recombined with a J_H region on chromosome 14. The two sites of breakage on chromosome 8 lie within 70 base pairs (bp) of one another. At each joining site, sequences homologous to the signal sequences thought to be recognized by the V-D-J recombinase⁹ were identified, as were N regions⁹. In B-cell chronic lymphocytic leukaemias (B-CLL) carrying the t(11; 14) chromosome translocation¹⁰ and in follicular lymphomas carrying the t(14; 18) translocation¹¹, the V-D-J recombinase is implicated in the mechanism of chromosomal translocations¹¹. We speculate that the same enzymatic mechanism is responsible for the t(8; 14) translocations in African Burkitt's lymphoma and pre-B cell ALL.

We have previously described⁸ the establishment of a cell line (380) from a patient with a pre-B-ALL (L2 in the French-American-British classification) carrying the t(8; 14) and t(14; 18) translocations. Using probes specific to the C_μ region we were able to show that both translocations involve the heavy chain locus, and to clone the breakpoint for the t(8; 14) translocation, which we denoted λ 380j9 (ref. 12, and Fig. 1). To establish which sequences from λ 380j9 derive from chromosome 8, we subcloned the 5' most SstI fragment (p380j9 0.8SS), and used it to probe a Southern blot of DNA from rodent $\times$ human somatic cell hybrids containing only chromosomes 8, 14 or the 14q⁺ chromosome from the African Burkitt's lymphoma cell line P3HR-1 (Fig. 2a). Hybridization occurred only in lanes containing DNA of human chromosome 8, or of the P3HR-1 14q⁺ chromosome. Hence this sequence derives solely from chromosome 8, and this portion of chromosome 8 is translocated to chromosome 14 in P3HR-1.

Using this chromosome 8 probe, we screened DNAs from normal and neoplastic tissues in order to detect rearrangements following translocations (Fig. 2b). Rearrangement was detected in the *Bam*H1 digest of P3HR-1 DNA, as well as the *Hind*III digest of Daudi African Burkitt's lymphoma DNA. These results indicate that the breakpoints in the 380 and P3HR-1 cell lines lie within the same *Bam*H1 fragment, while the Daudi breakpoint lies on a *Hind*III fragment nearby.

We attempted to clone the chromosome breakpoints in these cell lines, and to compare them to the 380 breakpoint clone (ref.

Received 17 February; accepted 24 September 1986.

1. Nishizuka, Y. *Nature* **308**, 693-698 (1984).
2. Fein, A. & Lisman, J. *Science* **187**, 1094-1096 (1975).
3. Lynch, G., Larson, J., Kelso, S., Barrionuevo, G. & Schottler, F. *Nature* **305**, 719-721 (1983).
4. Ekerot, C. F. & Oscarsson, O. *J. J. Physiol., Lond.* **318**, 207-221 (1981).
5. Kumamoto, E. & Kuba, K. *Nature* **305**, 345-346 (1983).
6. Dascal, N., Gillo, B. & Lass, Y. *J. Physiol., Lond.* **366**, 299-313 (1985).
7. Marty, A., Tan, Y. P. & Trautmann, A. *J. Physiol., Lond.* **357**, 293-325 (1984).
8. Rasmussen, H. & Barrett, P. Q. *Physiol. Rev.* **64**, 938-984 (1984).

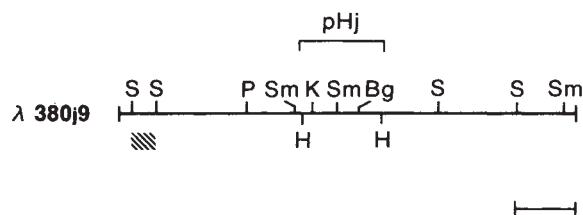


Fig. 1 Restriction map of the 14q⁺ breakpoint clone from the 380 B-ALL cell line, λ380j9. The region of the clone homologous to the joining segment probe pHj (ref. 12) is designated by a bracket. The hatched box denotes the *Sst*I fragment subcloned as probe p380j9 0.8SS. Indicated restriction sites are S, *Sst*I; P, *Pst*I; Sm, *Sma*I; H, *Hind*III; K, *Kpn*I and Bg, *Bgl*II. Scale bar 2 kb.

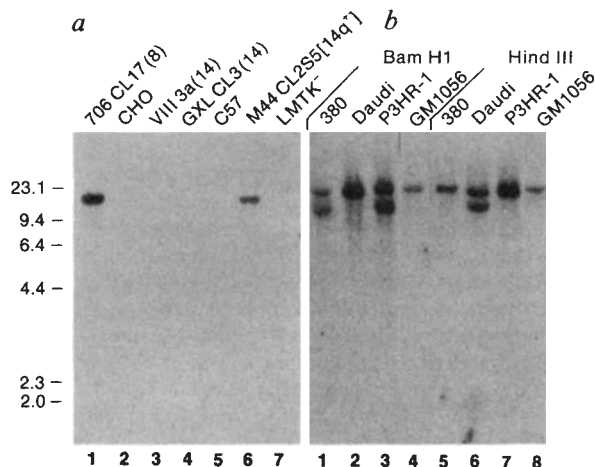


Fig. 2 *a*, Southern blot hybridization of *Hind*III digested DNAs from human × rodent somatic cell hybrids containing only human chromosome 8, 14 or the P3HR-1 14q⁺, probed with the p380j9 0.8SS probe. Lane 1, 706 Cl 17, a hybrid between a normal human lymphoblast and a Chinese hamster ovary (CHO) fibroblast containing human chromosome 8 but not 14; lane 2, CHO control; lane 3, GM 637 × C57 VIII 3a, a hybrid between a simian virus 40-transformed human fibroblast and a mouse parent, containing human chromosome 14 but not 8; lane 4, GM 119 × LMTK⁻ GXL C13, a hybrid between a human fibroblast and mouse LMTK⁻, also containing human 14 but not 8; lane 5, C56 control; lane 6, P3HR-1 × LMTK⁻ M44 Cl 2S5, a hybrid containing only the 14q⁺ chromosome from the African Burkitt's lymphoma cell line P3HR-1; lane 7, LMTK⁻ control. *b* Hybridization of 380, Daudi, P3HR-1 and GM 1056 genomic DNAs, with the p380j9 0.8SS probe. DNAs were digested with *Bam*HI (lanes 1–4) or *Hind*III (lanes 5–8). Lanes 1 and 5, 380 B-ALL cell line; lanes 2 and 6, Daudi Burkitt's lymphoma cell line; lanes 3 and 7, P3HR-1 Burkitt's lymphoma cell line; lanes 4 and 8, GM 1056 lymphoblastoid cell line.

Methods. Genomic DNA (10 µg) was digested with the indicated restriction enzyme, separated on a 0.7% agarose gel, and transferred to a nitrocellulose filter following denaturation and neutralization. The probe was ³²P-labelled by nick-translation. Hybridization was carried out in 4 × SSC, 2 × Denhardt's solution and 50% formamide at 37 °C for 24 h, and washing was to a final stringency of 2 × SSC at 65 °C. Blots were exposed for 12 h at -70 °C with an intensifying screen. The size of the λ *Hind*III digested marker DNA is in kb.

12 and Fig. 1). Probe p380j9 0.8SS was used to screen the 380, P3HR-1 and Daudi libraries. Two clones were obtained from the 380 library and one each from P3HR-1 and Daudi. These clones were subjected to analysis by restriction endonuclease digestion and Southern blotting. The resulting physical restriction maps are illustrated in Fig. 3. Several features are apparent. First, the clones are identical in the region containing the p380j9 0.8SS sequence derived from chromosome 8. Second, the Daudi (λD4) and 380 λ380-6B and λ380-8A clones overlap and

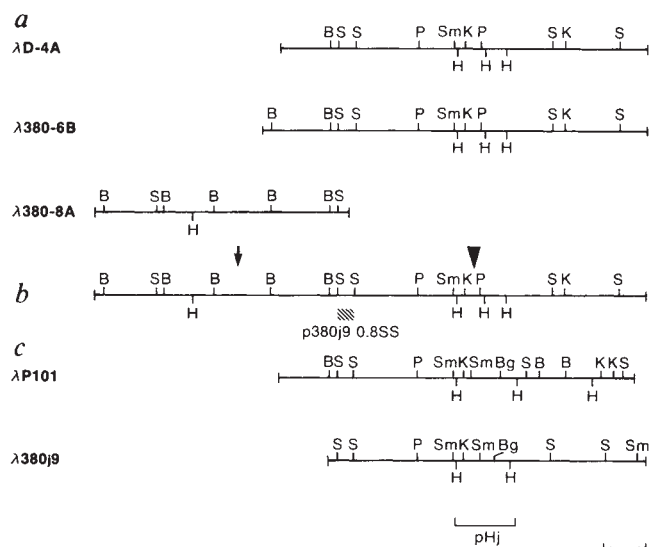


Fig. 3 *a*, Restriction maps of the normal chromosome 8 recombinant clones from the Daudi Burkitt's lymphoma (λD-4A) and the 380 B-ALL (λ380-6B and 380-8A) cell lines. The horizontal lines represent insert DNA from the respective clones. *b*, Restriction map of the normal chromosome 8 region deduced from *a*. Triangle, P3HR-1 and 380 breakpoints; arrow, putative Daudi breakpoint. *c*, Restriction map of the 14q⁺ breakpoint clone from the P3HR-1 Burkitt's lymphoma (λP101) and 380 (λ380j9) cell lines. The point at which the breakpoint clones diverge from the chromosome 8 map is indicated by the triangle in *b*. Cleavage sites of the restriction enzymes S, *Sst*I; P, *Pst*I; Sm, *Sma*I; H, *Hind*III; K, *Kpn*I and Bg, *Bgl*II. Scale bar, 2 kb. A representative genomic library was prepared from DNA of P3HR-1 and Daudi cells. Cloning procedures in the λ phage EMBL3A (ref. 34) were essentially as described¹². Libraries were screened with the p380j9 0.8 SS probe. Approximately 350,000 independent recombinant phages in the P3HR-1 library were screened, as well as approximately 800,000 phages in the Daudi library, and 350,000 clones in the previously constructed 380 genomic library¹². The P3HR-1 library was initially screened with the pHj joining segment probe¹². Six recombinant clones were obtained. These were secondarily screened with the chromosome 8-specific p380j9 0.8SS DNA probe (Fig. 1). Of the six initial positives, one clone, designated λP101, contained homologous sequences.

constitute the normal, uninterrupted chromosome 8 sequence. Third, the breakpoints of the 380 (λ380j9) and P3HR-1 (λP101) clones align almost precisely. Fourth, although the Daudi breakpoint clone was not obtained, genomic Southern blot data allow us to map the break to the middle of the 3.5 kb *Bam*HI fragment on chromosome 8 (indicated). Finally, comparison of this region with published restriction map¹³ of the sequences surrounding *c-myc* on chromosome 8 reveals no similarity: the breakpoints are situated 5' to the *Bam*HI site 15 kb upstream of the first *c-myc* exon.

Because the same site on chromosome 8 is involved in the translocation to chromosome 14 in P3HR-1 and 380, we sequenced this region for both the normal and breakpoint clones (Fig. 4). The 380 sequence is identical to that of the Daudi normal chromosome 8 allele but diverges six nucleotides 5' of the *Pst*I site. There follows a stretch of 22 nucleotides having homology to neither chromosome 8 nor chromosome 14 sequences. The next nucleotide corresponds to the twenty-first nucleotide of the heavy chain joining segment J6 (ref. 14), and the remainder of the 380 clone is homologous to this segment, with several single base substitutions. The P3HR-1 breakpoint is 63 bp 5' of the 380 breakpoint on chromosome 8. Upstream of this site, λP101 is homologous to λD4; 3' it is homologous to a region beginning 90 nucleotides 5' of J5. This result verifies the somatic cell hybrid data¹⁵ implicating the J_H segments in the P3HR-1 translocation.

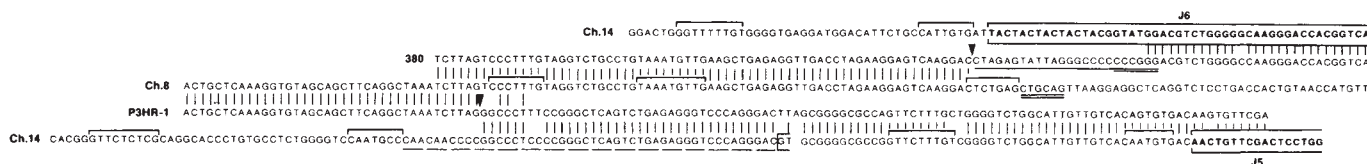


Fig. 4 Nucleotide sequence of the joining site between chromosome 8 and 14 in P3HR-1 and 380, and the corresponding normal chromosome 8 region. Triangles designate points at which the P3HR-1 and 380 sequences diverge from the chromosome 8 sequence; vertical lines indicate identical nucleotides. Bold sequences indicate the J5 and J6 heavy chain joining segments²³. The bracketed regions are conserved heptamer-nonamer signal sequences, and a putative N region is underlined at the 380 joining site. The underlined region of chromosome 14 comprises 16 codons and a splice signal (see text). DNA sequences were obtained using the Sanger dideoxy chain termination method³³ on M13³⁴ subclones from the *Kpn*I site (Fig. 3) through the chromosome 14 sequences. The *Pst*I site on λ D-4A, and the *Sma*I site on λ 380j9 (Fig. 3) are denoted by double underlines.

Both these sequences exhibit features suggesting that the recombinase mediating V-D-J joining was involved in these translocations. This enzyme is responsible for the rearrangement of the immunoglobulin locus during B-cell maturation⁹, and is thought to function similarly in T cells¹⁶. The putative enzyme is thought to recognize heptamer-nonamer signal sequences¹⁷ apposed to each recombining segment. Recombination only occurs when a signal sequence having a 12 base spacer is paired with one having a 23 base spacer¹⁸. Moreover, V-D-J recombination results in deletions, substitutions⁹, or in the addition of extra nucleotides (N regions), presumably by the action of terminal deoxynucleotide transferase¹⁹. The 380 joining site shows several of these points. The breakpoint on chromosome 8 occurs at a sequence homologous to the canonical heptamer (bracketed in Fig. 4). Although a satisfactory nonamer is lacking, a heptamer may be sufficient for recombinase function^{20,21}. The alignment of the J6 segment with the 380 sequence clearly demonstrates that the heptamer-nonamer on chromosome 14 are ideally positioned for a recombination event. Additionally, the stretch of extra bases corresponds to an N region, and includes a characteristic stretch of GCs⁹. At the P3HR-1 breakpoint, a heptamer is again positioned at the joining site on chromosome 8; twelve bases downstream is a nonamer. On chromosome 14, an almost perfect heptamer-nonamer is found, with a 23 bp spacer (shown in Fig. 4). Thus recombination according to the 12/23 rule could take place. Moreover, the 50 nucleotides 3' of the nonamer comprise 16 codons, including a stop codon, and a GT splice signal (Fig. 4). This sequence has the characteristics of a pseudo-joining segment, and could constitute an additional feature recognized by the recombinase. Imprecision is also observed at the breakpoint, as exemplified by the two base mismatch six bases 3' of the break. Lastly, both breakpoints exhibit single base changes which may be a consequence of somatic mutation following V-D-J recombination⁹. In both these translocations, then, nucleotide sequence analysis of the sites of chromosome-joining strongly suggests that the translocation occurred in a sequence-specific manner through the action of the recombinase.

Recombinase action can be related to the cell of origin of these tumours. Both originated in a pre-B cell²² or sIg⁺ cell^{20,21} capable of further immunoglobulin gene rearrangement. Taken together with similar findings in B-CLL¹⁰ and follicular lymphoma¹¹, these data suggest that tumour polyclonality²³ should be interpreted as continued gene rearrangement following malignant transformation. Moreover, in most sporadic Burkitt's lymphoma cases the chromosome breakpoints are to the first exon of *c-myc*²⁴⁻²⁹, and involve the μ , γ , or α switch regions²⁴⁻²⁹. However, in most African Burkitt's lymphomas, rearrangements of the *c-myc* locus cannot be detected by Southern blot hybridization^{4,30}, and translocations into the V_H (ref. 2) or J_H (ref. 15) region have been found. We and others have previously shown that B cells from sporadic Burkitt's lymphomas are more differentiated than those from African Burkitt's lymphomas³⁰⁻³². Our results support this hypothesis and suggest that in the African Burkitt's lymphoma P3HR-1, and in the 380 B-ALL, the translocation occurred at the pre-B cell stage of differenti-

ation during heavy chain gene rearrangement. The recombinase joined two separated segments of DNA on different chromosomes instead of two segments on the same chromosome. Thus in the endemic (African) Burkitt's lymphomas the translocations probably result from mistakes in V-D-J joining while in the sporadic cases it is possible that they result from mistakes during heavy chain class switching in more differentiated B cells.

Finally, comparison of the region of chromosome 8 involved in the translocations with the region surrounding *c-myc*¹³ shows no overlap. Calculations from the extent of the *c-myc* restriction map¹³, which covers ~20 kb 5' of the first exon, and the extent of our clones, which cover at least 15 kb away from the chromosome 14 junction, lead us to conclude that *c-myc* is at least 35 kb distant from the breakpoint in these instances. The precedent for transcriptional control acting over such distances comes from the IgH enhancer, which has been demonstrated to act over at least 17.5 kb³³. In each of these translocations, *c-myc* is situated 5' of the enhancer located between J_H and S _{μ} . This would appear to be the logical explanation for the deregulation of *c-myc* secondary to the t(8;14) chromosome translocation.

We thank Dr Kay Huebner for supplying the somatic cell hybrid DNAs, Dr Masaharu Isobe for many useful comments, Ms Debbie Jampetti for technical assistance and Mrs K. Reiner-smann for preparation of this manuscript. F.G.H. is a trainee of the Medical Scientist Training Program, NIH 5 T32 GM 07170. Supported by The Outstanding Investigator Award from the National Cancer Institute (C.M.C.).

Received 26 June; accepted 15 September 1986.

- Croce, C. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 3416-3419 (1979).
- Erikson, J., Finan, J., Nowell, P. C. & Croce, C. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5611-5616 (1982).
- Dalla Favera, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7824-7827 (1982).
- Dalla Favera, R., Martinotti, S., Gallo, R. C., Erikson, J. & Croce, C. M. *Science* **219**, 963-967 (1983).
- Taub, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7837-7841 (1982).
- Manolov, G. & Manolova, Y. *Nature* **237**, 33-34 (1972).
- Williams, D. L. *et al. Cell* **36**, 101-109 (1984).
- Pegoraro, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 7166-7170 (1984).
- Tonegawa, S. *Nature* **302**, 575-581 (1983).
- Tsujimoto, Y. *et al. Nature* **315**, 340-343 (1985).
- Tsujimoto, Y., Gorham, J., Crossman, J., Jaffe, E. & Croce, C. M. *Science* **226**, 1390-1393 (1985).
- Tsujimoto, Y., Finger, L. R., Yunis, J., Nowell, P. C. & Croce, C. M. *Science* **226**, 1097-1099 (1984).
- Bernard, O., Cory, S., Gerondakis, S., Webb, E. & Adams, J. M. *EMBO J.* **2**, 2375-2383 (1983).
- Ravetch, J. V., Sienbenlist, U., Korsmeyer, S., Waldmann, T. & Leder, P. *Cell* **27**, 583-591 (1981).
- Erikson, J., ar-Rushdi, A., Drwina, H., Nowell, P. C. & Croce, C. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 820-824 (1983).
- Yancopoulos, G. D. *et al. Cell* **44**, 251-259 (1986).
- Sakano, H., Huppi, K., Henrich, G. & Tonegawa, S. *Nature* **280**, 288-294 (1979).
- Sakano, H., Kurosawa, Y., Weigert, M. & Tonegawa, S. *Nature* **290**, 562-565 (1981).
- Desiderio, S. V. *et al. Nature* **311**, 752-755 (1984).
- Reth, M., Gehrmann, P., Petrac, E. & Wiese, P. *Nature* **322**, 840-842 (1986).
- Kleinfeld, R. *et al. Nature* **322**, 843-846 (1986).
- Korsmeyer, S. J. *et al. J. clin. Invest.* **71**, 301-313 (1983).
- Cleary, M. L. & Sklar, J. *Lancet* **i**, 1984, 489 (1984).
- Hamlyn, P. H. & Rabbitts, T. H. *Nature* **304**, 135-139 (1983).
- Gelmann, E. P., Psalidopoulos, M. C., Papas, T. S. & Dalla Favera, R. *Nature* **306**, 799-803 (1983).
- Wiman, K. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 6798-6802 (1984).
- Hamlyn, P. H. & Rabbitts, T. H. *Nature* **304**, 135-139 (1983).
- Gelmann, E. P., Psalidopoulos, M. C., Papas, T. S. & Dalla Favera, R. *Nature* **306**, 799-803 (1983).

26. Wiman, K. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 6798-6802 (1984).
27. Hamlyn, P. H. & Rabbitts, T. H. *Nature* **304**, 135-139 (1983).
28. Rabbitts, T. H., Hamlyn, P. H. & Baer, R. *Nature* **306**, 760-765 (1983).
29. Showe, L. C. *et al. Molec. cell. Biol.* **5**, 501-509 (1985).
30. Pelicci, P. G., Knowles, D. M., Magrath, I. & Dalla Favera, R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2984-2988 (1986).
31. Benjamin, D. *et al. J. Immun.* **129**, 1336-1342 (1982).
32. Croce, C. M. & Nowell, P. C. *Adv. Immun.* **38**, 245-274 (1986).
33. Wang, X. H. & Calame, K. *Cell* **43**, 659-665 (1985).
34. Murray, N. E., Hendrix, R. W., Roberts, J. W., Stahl, F., Weinberg, R. A. (eds) *Lambda II* (Cold Spring Harbor Laboratory, New York, 1983).
35. Sanger, F., Nicklen, S. & Coulson, A. E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
36. Messing, J. *Meth. Enzym.* **101**, 20-79 (1983).

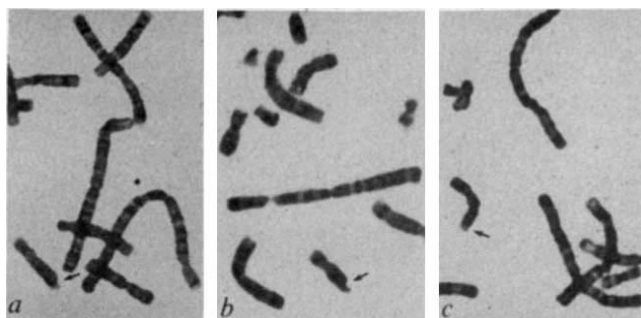


Fig. 1 Partial G-banded metaphases showing fragile X expression in hybrid cells from normal control males N1 (a), N2 (b), and a chimpanzee (c).

Implications of fragile X expression in normal males for the nature of the mutation

David H. Ledbetter*, Susan A. Ledbetter* & Robert L. Nussbaum†

* Kleberg Cytogenetics Laboratory, Institute for Molecular Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA

† Howard Hughes Medical Institute and Departments of Human Genetics and Pediatrics, University of Pennsylvania School of Medicine and Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA

The fragile site at Xq27, associated with a common form of X-linked mental retardation (XLMR)¹, is expressed in a variable proportion of the peripheral lymphocytes of affected males when the cells are cultured under thymidylate stress (Td stress) produced by folate or thymidylate deprivation. Some clinically normal males—transmitting males—are known to carry and transmit the fragile X mutation and yet show no cytogenetic expression in lymphocytes^{2,3}. Normal males with no family history of X-linked mental retardation express the site only rarely. When the fragile X chromosome from affected males is isolated in a rodent genetic background by somatic cell hybridization, the level of expression is similar to that seen in lymphocytes under Td stress⁴⁻⁶. Here we show that X chromosomes from two transmitting males and two normal control males, all of which were fragile X negative in lymphocytes or lymphoblasts, could be made to express the fragile site in hybrids, although at levels that were below those seen in hybrids from affected males. Furthermore, transmitting males could be differentiated from normal males by their significantly higher expression rates when hybrids were exposed to caffeine⁷ before cytogenetic harvest. One male chimpanzee also showed low level expression in hybrid cells. These data suggest that (1) the hybrid system lowers the threshold for fragile X expression, (2) a fragile site at Xq27 may be present on all human and chimpanzee X chromosomes and constitutes a previously unrecognized common fragile site and (3) the hybrid system with caffeine post-treatment can distinguish between the common Xq27 fragile site of control males, the occult mutant fragile site of a transmitting male, and the fully expressed fragile site of an affected male with XLMR. Thus the mutation producing XLMR may represent a multi-step alteration of a naturally occurring DNA sequence producing a continuum of cytogenetic expression and a threshold for clinical manifestation.

We and others have demonstrated that rodent cells (hamster or mouse) are permissive for fragile X expression when somatic cell hybrids are created containing the X chromosome of affected males^{4-6,8}. This approach allows the isolation of different fragile X mutations on a homogeneous, xenogeneic background. Hybrid expression levels are similar to those observed in the human parental cells, suggesting that expression is an intrinsic property of the site and is not influenced by the presence or absence of human autosomal loci^{4,6}. We have reported pre-

liminary hybrid studies of a clinically normal, transmitting male who showed no expression in lymphocytes or lymphoblasts in multiple laboratories^{6,7}. Low frequency expression (2%) was observed in hybrids and was enhanced to 12% by caffeine. This preliminary observation suggested that the hybrid system lowers the threshold for induction or detection of cytogenetic expression in individuals who carry the mutation but show no cytogenetic expression in routine lymphocyte studies. We have analysed additional hybrid clones from this transmitting male and have extended our studies to include a second unrelated transmitting male, two normal control males with no family history of X-linked mental retardation, and one male chimpanzee.

One to five hybrid clones from each individual were isolated, characterized, and used for fragile site induction studies. The clones studied, parental cell types and their chromosome constitution are shown in Table 1. The number of human chromosomes present varied from 1 (the X only) to 15, with an average of six.

Fragile X expression frequencies in parental cells and hybrid cells are shown in Table 2. The two affected males showed lymphoblast expression frequencies of 29% (GM4025) and 40% (DT) over several experiments. Under conditions of Td stress alone, hybrid expression was 27% and 23%, respectively. These levels were increased to 43% and 30% by caffeine post-treatment. Overall the hybrid expression rates were similar to those found in the parental cells.

In the two transmitting males, no expression was observed in 200 (MGL66) or 100 (LC) lymphoblast mitoses. This is consistent with previous studies, in which lymphocytes and lymphoblasts have been repeatedly negative (refs 9, 10; P. A. Jacobs, unpublished; W. T. Brown, unpublished). Under Td stress alone, hybrid cells from each transmitting male showed 2% expression (22/900 and 9/400) over multiple experiments. With caffeine, expression increased in the hybrids from both transmitting males to 12% (62/500 and 12/100). This significant enhancement by caffeine is similar to that previously observed in a low-expressing carrier female^{6,7}.

Several investigators have proposed that the occurrence of clinically normal, transmitting males may be due to the presence of one or more autosomal modifying loci which suppress the major gene on the X chromosome^{2,11}. Segregation data argue against this hypothesis¹², but no previous experiments have addressed the issue. Segregation of the proposed autosomal suppressor by somatic cell hybridization might be expected to produce expression levels equivalent to those of affected males. The significantly lower expression rates observed in six independent clones from two unrelated transmitting males suggest an alternative explanation, that the hybrid system lowers the threshold for detection of low level fragile site expression in transmitting males with an occult or incomplete mutation.

A total of 100 mitoses from peripheral lymphocytes were scored from each of two normal control males with no family

Table 1 Characterization of somatic cell hybrids used for fragile X expression

Human/chimpanzee parent	Hybrid clone	Hamster parent	Human/chimpanzee chromosomes	Ref.*
GM4025	4.12	RJK88	X	4, 6
	PHL3	88.13	X, 3	5
DT	DT10	88.13	X, 3, 4, 7, 10, 16, 17, 20, 21, 22	—
	DT6	88.13	X, 1, 4, 5, 6, 7, 9, 11, 12, 14, 16, 18, 21, 22, ?Y	—
MGL66	MGL66-1	RJK88	X, 13	6
	MGL66-2	RJK88	X, 4, 6, 21, 22, der(9)	—
	MGL66-3	RJK88	X, 4, 11, 12, 15, 20, 21, der(9)	—
	MGL66-4	RJK88	X, 6	6
	MGL66-5	RJK88	X, 12, 14, 15, 18, 21, Y	—
LC	LC1A	RJK88	X, 11, 21, ?Y, ?18	—
N1	RJK739	RJK88	X	—
N2	DHL3A	RJK88	X, 3, 3, 7, 8, 10, 11, 12, 13, 14, 18, 20, 21, 22	—
AG6939A	B4	RJK88	X + approximately 32 autosomes	—

Parental cell types used for fragile X expression and somatic cell fusion were either lymphoblasts, lymphocytes or fibroblasts. GM4025 is a lymphoblast line from an affected fragile X male obtained from the Human Genetic Mutant Cell Repository. DT is a lymphoblast line produced by Epstein-Barr-virus transformation of an unrelated affected male with a 40% lymphocyte expression rate¹. MGL66 is a lymphoblast line provided by Dr P. A. Jacobs from a transmitting male who showed no expression in lymphocytes⁹. LC is a lymphoblast line provided by Dr W. T. Brown from a transmitting male with no expression in lymphocytes¹⁰. Peripheral blood was obtained from two clinically normal control males (N1, N2) with no family history of mental retardation for fragile X expression and for fusion. A fibroblast cell line on a male chimpanzee (*Pan troglodytes*) was obtained from the Human Genetic Mutant Cell Repository (AG6939A). Hybrids were constructed as previously described⁴ from parental cells and RJK88, an HPRT⁻ Chinese hamster V79 derivative or 88.13, a thymidylate synthase deficient (TS⁻) mutant of RJK88⁵. Hybrids were maintained in Dulbecco's modified Eagle's medium + 10% FCS with hypoxanthine (10^{-4} M) and azaserine (5×10^{-4} M). The chromosome composition of each hybrid was determined by trypsin G-banding analysis.

* Data are from this work except where stated.

Table 2 Per cent fragile X expression in parental and hybrid cells from affected males, transmitting males, normal control males and a chimpanzee male

Cell type/treatment	Affected GM4025	DT	Transmitting MGL66	LC	Normal control N1	N2	Chimpanzee AG6939A
Parental—							
Td stress	29 (58/200)	40 (60/150)	0 (0/200)	0 (0/100)	0 (0/200)	0 (0/200)	—
Hybrids							
Td stress	27 (126/468)	23 (34/150)	2 (22/900)	2 (9/400)	2 (5/250)	2 (3/200)	0.2 (1/400)
Caffeine	43 (152/350)	30 (30/100)	12 (62/500)	12 (12/100)	4 (8/200)	5 (9/187)	1.6 (4/250)

The number of fragile X positive cells observed over mitoses scored is given in parentheses. Each condition represents at least two experiments. For fragile X induction, lymphoblasts and lymphocytes were cultured in medium RPMI 1640 + 10% FCS and treated with FUDr (5×10^{-8} M and 1×10^{-7} M) or excess thymidine (1–2 mM) for the last 24 hours before standard cytogenetic harvest. For hybrids, induction by thymidylate (Td) stress alone was by treatment with FUDr (10^{-9} M– 10^{-7} M) or excess thymidine (1–2 mM) in most hybrids, and by thymidine deprivation in the TS⁻ hybrids⁵ for 12–16 hours. Caffeine post-treatment was used to enhance expression by addition of 2 mM caffeine for the last 1–2 hours before harvest⁷. All slides were trypsin G-banded and at least 50 cells per condition scored for fragile X expression whenever possible. Scoring of the lymphocytes and hybrids from the two normal control males was done blindly in the sense that the technicians doing the analysis were unaware of the nature of the experiment and the clinical status of the individuals whose cells they were scoring.

history of X-linked mental retardation. No fragile X chromosomes were observed. However, in blind experiments, 2% expression (5/250 and 3/200) was consistently observed in hybrid cells under Td stress alone. This level is identical to that observed in the clinically normal transmitting males. Caffeine post-treatment increased expression to 4% (8/200) and 5% (9/187), significantly less than the 12% level of transmitting males ($\chi^2 = 16.8$, $P < 0.01$). The total of 25 fragile X positive cells observed in the two normal individuals (examples in Fig. 1) is much higher than expected by random chromosome breakage. Our findings suggest that previous reports of single cell occurrences of fragile X cells in normal controls (reviewed in ref. 13) are not artefacts, but represent low level expression of a common fragile site at Xq27 in normal individuals.

Common fragile sites are present in both homologues of a chromosome in most or all individuals, whereas rare fragile sites are usually found in just one homologue and occur in less than 1% of the population¹⁴. The relationship between rare and common fragile sites is unclear. Their mechanism of induction is similar, though not identical¹⁵. Interestingly, 6 of 18 rare fragile sites occur in chromosome bands which also contain a

common fragile site¹⁴. This has led to speculation that a rare fragile site may represent an alteration (mutation) in a pre-existing common fragile site^{16,17}. Yunis and Soreng¹⁶ demonstrated the presence of common fragile sites in chimpanzee and gorilla in sites homologous to those of man. If Xq27 is a common fragile site in humans, it should also be present in these species. We therefore extended our hybrid studies to a male chimpanzee. No fragile X induction studies were attempted on the parental fibroblast cells, as fragile X expression is generally poor in fibroblasts compared to lymphocytes or lymphoblasts even in affected fragile X males. In hybrid cells under Td stress alone, one fragile X cell was observed out of 400 mitoses scored. With caffeine post-treatment, 4 positive cells out of 250 were seen (Fig. 1c). Although this frequency of expression is lower than that seen in normal human males ($\chi^2 = 4.98$, $P < 0.05$), our impression is that this may be due to the relatively poor growth of this hybrid clone. Despite the lower frequency, the presence of a fragile site at Xq27 in a chimpanzee is consistent with the existence of a common fragile site whose origin is ancestral to the human-chimpanzee divergence.

Use of caffeine to enhance fragile X expression in hybrid cells

was crucial to the discrimination between transmitting and normal males. Yunis and Soreng¹⁶ first showed that caffeine, which has synergistic effects with a variety of DNA damaging agents, enhances expression of common fragile sites and fragile X when used in conjunction with fluorodeoxyuridine (FUDR). Although we⁷ and others¹⁸⁻¹⁹ have failed to replicate caffeine's enhancing effect on fragile X expression in lymphocytes or lymphoblasts, it has consistently increased expression in hybrid cells^{6,7}. This difference may be due to the fact that caffeine's effect on fragile X expression appears to be indirect: it inhibits the mitotic delay usually associated with DNA damage, thereby reducing the time available for DNA replication, repair and chromosome condensation⁷. The effect may vary among cell types with different cell cycle characteristics, particularly different lengths of G₂.

The high prevalence and unusual segregation patterns observed in fragile X syndrome suggest that it is not a classical X-linked mutation^{3,12,20}. The single most unusual characteristic of its inheritance is its variable penetrance within pedigrees. For example, the penetrance of mental retardation is significantly lower in offspring of mothers of transmitting males than in offspring of daughters of transmitting males¹². This observation has been referred to as the 'Sherman paradox'²¹ and must be explained by any model addressing inheritance of fragile X. The unique ability of the hybrid system to induce fragile X expression in transmitting and normal males and to demonstrate a continuum of cytogenetic expression (low in normal males, higher in transmitting males, highest in affected males) argues strongly against the involvement of autosomal modifying loci in this disorder. Instead, the presence of a common fragile site at Xq27 is consistent with the multi-step mutation models of Pembrey *et al.*²⁰ and Nussbaum *et al.*²².

We speculate that recombination between two normal X chromosomes at the Xq27 common fragile site region may be the mechanism for an initiation event, producing a clinically normal transmitting male or carrier female. Segregation data demonstrates that progression to the full XLMR mutation occurs only in females²⁰ (who have two X chromosomes), also suggesting that meiotic recombination is involved. These recombination events may produce variation in the length or copy number of a DNA sequence at Xq27, which at some threshold level disrupts the function of a neighbouring gene and produces XLMR. Quantitative variation in the size of the fragile site DNA region may also explain the variable recombination distances between polymorphic DNA markers observed among fragile X families²³. Thus, our data suggest that a normal DNA sequence at Xq27 may be altered to produce a continuous quantitative variation in fragile site DNA resulting in varying degrees of cytogenetic expression and a threshold for clinical manifestation of the mutation.

We thank Drs P. A. Jacobs and W. T. Brown for providing lymphoblastoid cell lines on transmitting males. This work was supported by NIH RO1 HD20227, the Robert J. Kleberg, Jr Center for Human Genetics, and the Howard Hughes Medical Institute. R.L.N. is an Associate Investigator in the Howard Hughes Medical Institute.

13. Jenkins, E. C., Brown, W. T., Brooks, J., Duncan, C. J., Sanz, M. M. *et al.* *Exp Cell Biol.* **54**, 40-48 (1986).
14. Berger, R., Bloomfield, C. D. & Sutherland, G. R. *Cytogenet. Cell Genet.* **40**, 490-535 (1985).
15. Glover, T. W. in *Fragile Sites on Human Chromosomes* (eds Sutherland, G. R., Hecht, F., Mulley, J. C., Glover, T. W., Hecht, B. K.) 80-94 (Oxford University Press, New York, 1985).
16. Yunis, J. J. & Soreng, A. L. *Science* **226**, 1199-1204 (1984).
17. Daniel, A. *Am. J. med. Genet.* **23**, 419-427 (1986).
18. Glover, T. W., Coyle-Morris, J. & Morgan, R. *Cancer Genet. Cytogenet.* **19**, 141-150 (1986).
19. Abruzzo, M. A., Pettay, D., Mayer, M. & Jacobs, P. A. *Hum. Genet.* **73**, 20-22 (1986).
20. Pembrey, M. E., Winter, R. M. & Davies, K. E. *Am. J. med. Genet.* **21**, 709-718 (1985).
21. Opitz, J. M. *Am. J. med. Genet.* **23**, 1-10 (1986).
22. Nussbaum, R. L., Airhart, S. D. & Ledbetter, D. H. *Am. J. med. Genet.* **23**, 715-722 (1986).
23. Brown, W. T., Gross, A. C., Chan, C. B. & Jenkins, E. C. *Am. J. med. Genet.* **23**, 643-664 (1986).

Analysis of enzymatically amplified β -globin and HLA-DQ α DNA with allele-specific oligonucleotide probes

Randall K. Saiki, Teodorica L. Bugawan,
Glenn T. Horn, Kary B. Mullis & Henry A. Erlich

Department of Human Genetics, Cetus Corporation,
1400 Fifty-Third Street, Emeryville, California 94608, USA

Allelic sequence variation has been analysed by synthetic oligonucleotide hybridization probes which can detect single base substitutions in human genomic DNA¹⁻³. An allele-specific oligonucleotide (ASO) will only anneal to sequences that match it perfectly, a single mismatch being sufficient to prevent hybridization under appropriate conditions⁴. To improve the sensitivity, specificity and simplicity of this approach, we used the polymerase chain reaction (PCR) procedure⁵ to enzymatically amplify a specific segment of the β -globin or HLA-DQ α gene in human genomic DNA before hybridization with ASOs. This *in vitro* amplification method, which produces a >10⁵-fold increase in the amount of target sequence, permits the analysis of allelic variation with as little as 1 ng of genomic DNA and the use of a simple 'dot blot' for probe hybridization. As a further simplification, PCR amplification has been performed directly on crude cell lysates, eliminating the need for DNA purification.

³²P-labelled ASO probes have been used in the analysis of a number of genetic diseases caused by single base mutations, including sickle cell anaemia¹, α -1 anti-trypsin deficiency² and β -thalassaemia³. Although the ASO method is a powerful diagnostic tool, it requires 5 to 10 μ g of genomic DNA, gel-purified high specific activity probes and several days of autoradiographic exposure. The short oligomer probes (usually 19 bases long) also hybridize to other genomic sequences so that restriction endonuclease digestion and gel electrophoresis are required to separate the target sequence from the bulk of the genomic DNA. In order to develop a simple, sensitive method, we chose the sickle cell anaemia and haemoglobin C mutations in the sixth codon of the β -globin gene⁶ as a model system for genetic diagnosis. We used ASOs specific for the normal (β^A), sickle cell (β^S), and haemoglobin C (β^C) sequences as probes to detect these alleles in PCR-amplified genomic samples. Figure 1 shows the sequence of the PCR primers (PC03 and PC04), the ASO probes (19A, 19S and 19C), and their relationship to the β -globin gene. The PCR primers, both 20 bases in length, anneal to opposite strands of the genomic DNA flanking the target sequence⁷. These primers are positioned so that the polymerase-catalysed, template-directed extension product of one primer can serve as a template strand for the other. In this configuration, repeated cycles of denaturation, primer annealing, and primer extension, result in the exponential accumulation of a 110-base pair (bp) β -globin fragment, with a length defined by the distance between the 5' ends of the PCR primers. For example, 20 cycles of PCR amplification with PC03 and PC04 as primers produce a 220,000-fold increase in the amount of β -globin target

Received 30 May; accepted 26 August 1986.

1. Sutherland, G. R., Hecht, F., Mulley, J. C., Glover, T. W. & Hecht, B. K. (eds) *Fragile Sites on Human Chromosomes*, 304 (Oxford University Press, New York, 1985).
2. Froster-Iskenius, U., McGillivray, B. C., Dill, F. J., Hall, J. G. & Herbst, D. S. *Am. J. med. Genet.* **232**, 619-632 (1986).
3. Howard-Peebles, P. N. & Friedman, J. M. *Am. J. hum. Genet.* **37**, 956-964 (1985).
4. Nussbaum, R. L., Airhart, S. D. & Ledbetter, D. H. *Hum. Genet.* **64**, 148-150 (1983).
5. Nussbaum, R. L., Walmsley, R. M., Lesko, J. G., Airhart, S. D. & Ledbetter, D. H. *Am. J. hum. Genet.* **37**, 1192-1205 (1985).
6. Ledbetter, D. H., Airhart, S. D. & Nussbaum, R. L. *Am. J. med. Genet.* **23**, 429-444 (1986).
7. Ledbetter, D. H., Airhart, S. D. & Nussbaum, R. L. *Am. J. med. Genet.* **23**, 445-456 (1986).
8. Warren, S. T. & Davidson, R. L. *Somat. Cell molec. Genet.* **10**, 409-413 (1984).
9. Rhoads, F. A., Oglesby, A. C., Mayer, M. & Jacobs, P. A. *Am. J. med. Genet.* **12**, 205-217 (1982).
10. Brown, W. T., Gross, A. C., Chan, C. B. & Jenkins, E. C. *Hum. Genet.* **71**, 11-18 (1985).
11. Steinbach, P. *Hum. Genet.* **72**, 248-252 (1986).
12. Sherman, S. L., Jacobs, P. A., Morton, N. E., Froster-Iskenius, U. & Howard-Peebles, P. N. *Hum. Genet.* **69**, 289-299 (1985).

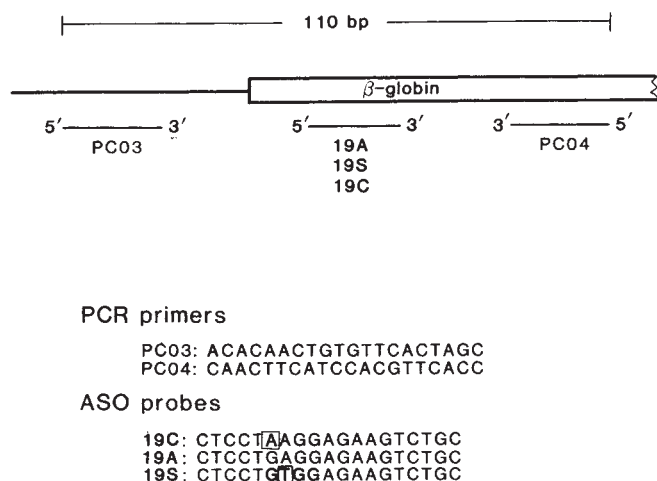


Fig. 1 Sequence of the PCR primers and ASO probes and their relationship to the β -globin gene. The two PCR primers, PC03 and PC04, are 20 bases long and are complementary to the (-) and (+) strands, respectively. The three ASO probes, 19A, 19S, and 19C, are 19 bases long and span codons 4-17 of the gene⁸. 19A is complementary to the normal β -globin allele (β^A), 19S to the sickle cell allele (β^S), and 19C to the haemoglobin C allele (β^C). Boxes, single base mutations of β^S and β^C .

sequence⁷. The sequences of the 19-base ASO probes used here are identical to those described previously by Studencki *et al.*⁸.

DNA was extracted from six blood samples from individuals with every possible diploid combination of the β^A , β^S , and β^C alleles and from the cell line GM2064, which has a homozygous deletion of the β -globin gene⁹. The DNA (1 μ g from each sample) was subjected to 25 cycles of PCR amplification and 1/30th of the reaction product (33 ng) applied to a nylon filter as a 'dot blot'. Three replicate filters were prepared and each was hybridized with one of the three ³²P-labelled ASOs under stringent conditions. The autoradiogram (Fig. 2) clearly indicates that each ASO annealed only to those DNA samples which contained at least one copy of the β -globin allele to which the probe was perfectly matched. For example, the β^A -specific probe, 19A, hybridized only to samples AA ($\beta^A\beta^A$), AS ($\beta^A\beta^S$), and AC ($\beta^A\beta^C$). The high degree of specificity contributed by PCR amplification is further emphasized by the lack of any detectable hybridization of the ASO probes to the GM2064 deletion mutant. The frequency of the specific β -globin target in PCR amplified DNA has been estimated by the analysis of cloned amplification products to be approximately 1%, an enrichment of $>10^5$ over unamplified genomic DNA¹⁰. It is this substantial reduction of complexity that allows the application of 19-base probes in a dot blot format and also makes possible the use of even shorter ASO probes capable of allelic discrimination under less stringent conditions (data not shown).

To estimate the minimum amount of DNA detectable, eight serial dilutions containing 1 to 128 ng of genomic AA DNA were amplified by 25 cycles and half of each sample was applied to a dot blot, then hybridized with the 19A probe. Equivalent dilutions of the 25-cycle amplifications of AA and SS DNA prepared in the previous experiment were included as controls. The autoradiogram (Fig. 3) reveals signals in all AA samples; the overall efficiency of amplification with nanogram quantities of template DNA is comparable to that obtained with 1 μ g. Moreover, the subnanogram detection limit obtained with ³²P-labelled oligonucleotide probes suggests that the analysis of nanogram quantities of genomic DNA with nonradioactive ASO probes (for example, biotin-labelled) should be feasible. Since one nanogram is the amount of genomic DNA present in only 150 diploid cells (assuming a haploid genome size of 3×10^9 bp)

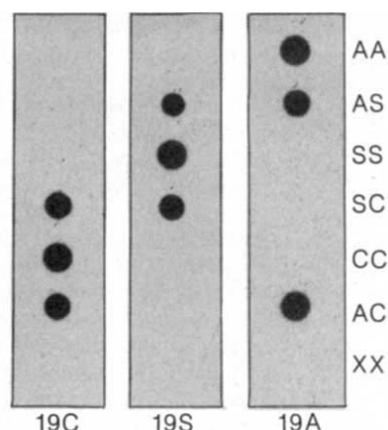


Fig. 2 Genotype analysis of PCR-amplified genomic DNA using ASO probes. DNA was extracted from the buffy coat fractions of clinical blood samples of known β -globin genotype¹¹ and 1 μ g portions of AA ($\beta^A\beta^A$), AS ($\beta^A\beta^S$), SS ($\beta^S\beta^S$), SC ($\beta^S\beta^C$), CC ($\beta^C\beta^C$), AC ($\beta^A\beta^C$), and GM2064 (deletion mutant, see text) DNA were subjected to 25 cycles of PCR amplification as described previously⁷. Briefly, the samples were dispensed in 100 μ l PCR reaction buffer (50 mM NaCl, 10 mM Tris, pH 7.6, 10 mM MgCl₂, 1 μ M PC03, 1 μ M PC04, 1.5 mM dATP, 1.5 mM dCTP, 1.5 mM dTTP, and 1.5 mM dGTP). After heating (10 min at 95 $^{\circ}$ C) to denature the DNA, the samples were transferred to a 30 $^{\circ}$ C heat block for 2 min to allow the primers to anneal to their target sequences. Primer extension was initiated by addition of one unit of Klenow fragment of *Escherichia coli* polymerase I (New England Biolabs) and incubated for 2 min at 30 $^{\circ}$ C. This cycle of heating, annealing, and extension was repeated 24 times; subsequent denaturations were by heating for 2 min. The final volume after 25 cycles was 150 μ l. One-thirtieth of each reaction, (5 μ l, containing 33 ng of the original genomic DNA), was adjusted to 0.4 N NaOH, 25 mM EDTA¹⁷ in a 200 μ l volume and applied to a Genatran-45 nylon filter (Plasco) with a Bio-Dot spotting apparatus (Bio-Rad). Three replicate filters were prepared. The ASO probes were phosphorylated at their 5'-termini with [γ -³²P]ATP and polynucleotide kinase and purified by spin dialysis¹¹. The specific activities of the probes were between 3.5 and 4.5 μ Ci pmol⁻¹. Each filter was prehybridized individually in 8 ml $5 \times$ SSPE, $5 \times$ Denhardt, 0.5% SDS for 30 min at 55 $^{\circ}$ C. Probe (1 pmol) was then added and hybridized at 55 $^{\circ}$ C for 1 h. Filters rinsed twice in $2 \times$ SSPE, 0.1% SDS at room temperature, followed by a high stringency wash in $5 \times$ SSPE, 0.1% SDS for 10 min at 55 $^{\circ}$ C (for 19C) or 60 $^{\circ}$ C (for 19A and 19S) and autoradiographed for 2.5 h at -80 $^{\circ}$ C with a single intensification screen (Dupont Lightning Plus).

we considered the possibility that a crude lysate from a small number of cells would contain sufficiently little cellular debris to allow the DNA to serve as template for the PCR amplification procedure. Eight dilutions containing 75 to 9,600 cells from each of the cell lines MOLT-4 ($\beta^A\beta^A$)¹¹, SC-1 ($\beta^S\beta^S$)¹¹, and GM2064 were heated at 95 $^{\circ}$ C to lyse the cells and subjected to 25 cycles of amplification without purification. Equivalent amounts of purified MOLT-4 DNA were included as controls. The amplified samples were divided into two dot blots and hybridized with the 19A and 19S probes. The autoradiogram (Fig. 4) shows a positive signal in an unfractionated cell lysate of only 75 cells. The efficiency of amplification appears to be only two- to three-fold less than with equivalent amounts of purified DNA. (MOLT-4 is reported to be 90% tetraploid (A. Green, personal communication) so that 150 cells would be expected to contain 2 ng of genomic DNA.) Both ASO probes retain allele specificity: 19A anneals only to MOLT-4 DNA and 19S only to SC-1 DNA. The variation from the expected two-fold reduction of signal intensities may reflect dilution artefacts associated with manipulating such small numbers of cells, and the significant reduction in signal intensity seen in both 9,600-cell samples may be due to inhibition of the amplification process by cellular

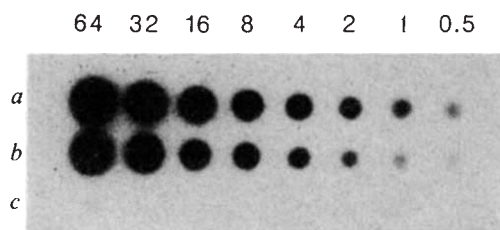


Fig. 3 PCR amplification of sub-microgram amounts of genomic DNA. *a*, Eight serial dilutions containing 128, 64, 32, 16, 8, 4, 2, and 1 ng of sample AA DNA were amplified by 25 cycles. One half of the amplification product (64 to 0.5 ng) was applied to a filter (row A). Controls consisted of equivalent dilutions of *b*, previously amplified AA and *c*, previously amplified SS DNA (Fig. 2). The filter was hybridized with the 19A probe and exposed for 6 h. See Fig. 2 (legend) for methods.

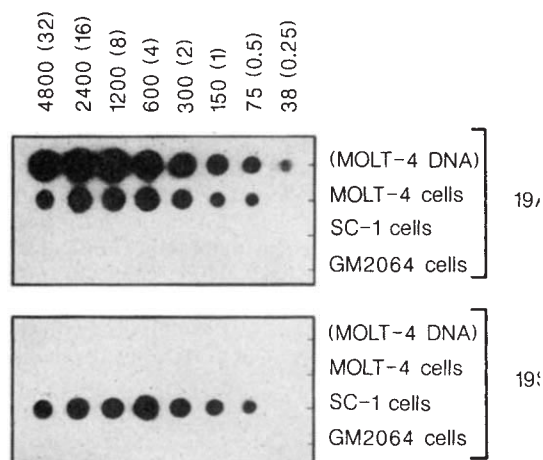


Fig. 4 Genotype analysis of PCR-amplified crude cell lysates using ASO probes. Cell concentrations of the cell lines MOLT-4 ($\beta^A\beta^A$), SC-1 ($\beta^S\beta^S$), and GM2064 (see text) were determined with a haemocytometer. Eight serial dilutions in PBS were prepared at 9.6×10^5 , 4.8×10^5 , 2.4×10^5 , 1.2×10^5 , 6.0×10^4 , 3.0×10^4 , 1.5×10^4 , and 7.5×10^3 cells per ml. Ten μ l of each dilution (containing 9,600, 4,800, 2,400, 1,200, 600, 300, 150 and 75 cells), were mixed with 35 μ l H_2O and heated for 10 min at $95^\circ C$ to lyse the cells. The lysates were added to 55 μ l, $1.8 \times$ PCR buffer (containing salts, primers, and dNTPs) and immediately subjected to 25 cycles of amplification. Equivalent amounts (64 to 0.5 ng) of purified MOLT-4 genomic DNA were also amplified. The PCR reaction products were divided onto two filters, probed with 19A and 19S and exposed for 6 h. Methods are as for Fig. 2 except that the amplification buffer included 10% DMSO (v/v), the annealing and extension reactions were performed at $37^\circ C$ instead of $30^\circ C$ ¹⁰ and the PCR amplification procedure was carried out in an automated device designed by the Cetus Instrumentation group.

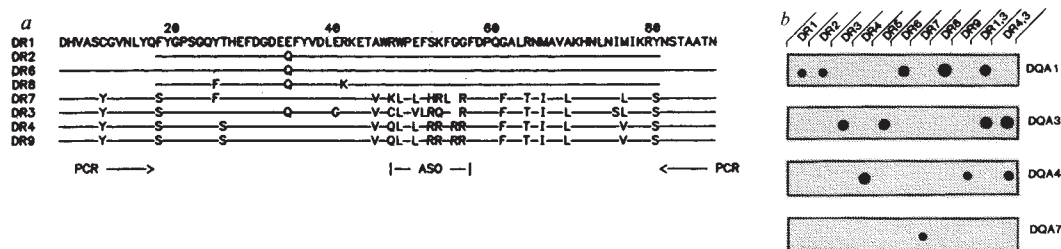
debris. Use of unfractionated lysate reduces the time required for analysis to <8 hours.

PCR amplification and ASO hybridization have also been used to analyse genetic polymorphism at the HLA-DQ α locus, which has localized allelic variability in the second exon (Fig. 5a). Using PCR primers complementary to conserved regions of this exon, a 242-bp fragment of DQ α was amplified from DNA samples isolated from nine homozygous typing cell lines (HTC) and from two heterozygous cell lines, then hybridized with four different ASO probes (Figure 5b). The sequence of these probes was based on published data¹²⁻¹⁵ as well as on our determinations of the sequences of complementary DNA, genomic and PCR-amplified clones (ref. 12 and Fig. 5, legend). Polymorphism in the HLA-DQ region, which consists of one α and one β chain locus, is detected serologically as the DQw1, DQw2, and DQw3 antigen series¹⁶. The DQA1 probe hybridizes selectively to amplified DNA from all cells typed as DQw1, which is associated with the DR1, DR2, and DRw6 serological types. As expected from the gene sequence, this probe also hybridizes to the DNA from the DR8 homozygous cell line TAB which is also DQw1. The DQA3 probe hybridizes to DNA from the DR3, DQw2 line AVL as well as to the unknown sequence in the DR5, DQw3 line JGL suggesting that these DQ α alleles are homologous. Since DQA3 did not hybridize to the DNA from the other DQw3 lines (DKB and KOZ), there must be

DQ α sequence heterogeneity within this specificity. Thus the allelic variation in the HLA-DQ α locus detected by this method appears to be greater than the polymorphic series defined by serological typing, and can also be more precisely localized. This approach is at present being applied to the analysis of polymorphism in the HLA-DQ β and DR β loci.

With the use of nonisotopic probes and PCR automation (see Figs 4 and 5), this procedure, combining *in vitro* target

Fig. 5 Detection of allelic variation in HLA-DQ α sequences. *a*, Allelic variation in the HLA-DQ α gene; *b*, detection of HLA-DQ α alleles. Protein sequences of the variable outer domain of the HLA-DQ α chain, derived from cloned cDNA and PCR-amplified¹⁰ sequences, are shown as differences from the sequence found in the DR1 (ref. 13) haplotype (upper line), and grouped to emphasize their similarities. The sequences from the DR1, DR2 (ref. 13) and DR6 (ref. 14) haplotypes (DQw1) are closely related, and those from the DR4 (ref. 14) and DR9 (ref. 16) haplotypes (DQw3) are identical. The region between positions 47 and 56 shows the greatest variation, and was therefore chosen as a target for four different ASO probes. In this segment, the DQ α sequences from the DR1, 2, 6, and 8 (ref. 13) haplotypes are identical and homologous to the DQA1 ASO (5'-AACCTCCAAATTTGCTGAAGTCAAGCCACCG-3'). The sequence from the DR7 (ref. 15) haplotype is homologous to the DQA7 ASO (TTCCACAGACTTAGATTG) and the DR3 sequence (ref. 14) to the DQA3 probe (TGTTTCCTGTTCTCAGAC). The sequences from the DR4 and DR9 haplotypes are identical and homologous to the DQA4 probe (TTCCGCAGATTAGAAGAT).



DNA was isolated¹¹ from nine homozygous typing cell lines and two heterozygous cell lines with HLA-DQ α types DR1 (LG2), DR2 (PGF), DR3 (AVL), DR4 (DKB), DR5 (JGL), DR6 (APD), DR7 (LBF), DR8 (TAB), DR9 (KOZ), DR1,3 (B302), DR3,4 (B303). Each DNA sample (1 μ g) was PCR-amplified for 28 cycles with the DQ α primers as described (Fig. 4). The degree of amplification was determined by hybridization to a DQ α cDNA probe labelled with ^{32}P by nick-translation, and equal amounts of product (2–10% of the final reaction volume) was denatured with two volumes of 0.5 M NaOH, 1.5 M NaCl. After neutralization with 0.5 M Tris, pH 7, 1.5 M NaCl to a final volume of 200 μ l, the DNA was spotted in four replicates on MSI Nylon 66 (Fischer Scientific) and fixed to the filters by UV irradiation for 4 min. The filters were prehybridized for 1 h at 50° in $6 \times$ SSPE, $10 \times$ Denhardt, 0.5% SDS, and hybridized overnight with the ASO probes (at 2×10^5 cpm ml⁻¹) labelled with ^{32}P as described (Fig. 2), washed for 10 min at 37° in $0.1 \times$ SSPE, 0.1% SDS, and autoradiographed for 90 min with an intensification screen.

amplification and ASO probes in a dot-blot format, promises to be a general and simple method for the detection of allelic variation. In addition, the ability to analyse genetic variation rapidly in minute amounts of purified DNA or cell lysates has important implications for a wide variety of genetic analyses.

We thank Corey Levenson, Lauri Goda and Dragan Spasic for the synthesis of oligonucleotides, Norm Arnheim for discussions, and Larry Johnson and Richard Leath for instrument and engineering support. We also thank John Sninsky for critical review, and Kathy Levenson, Eric Ladner, Tim Culp and Sharon Nilson for preparation of this manuscript.

Note added in proof: Recently, 50 ng samples of PCR-amplified genomic DNA have been analysed using a biotin-labelled ASO probe.

Received 30 May; accepted 5 September 1986.

1. Conner, B. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 278–272 (1983).
2. Kidd, V. J., Wallace, R. B., Itakura, K. & Woo, S. L. C. *Nature* **304**, 230–234 (1983).
3. Pirastu, M. *et al. New Engl. J. Med.* **309**, 284–287 (1983).
4. Wallace, R. B. *et al. Nucleic Acids Res.* **6**, 3543–3557 (1979).
5. Faloona, F. A. & Mullis, K. B. *Meth. Enzym.* (in the press).
6. Kan, Y. W. & Dozy, A. M. *Science* **209**, 388–391 (1980).
7. Saiki, R. K. *et al. Science* **230**, 1350–1354 (1985).
8. Studnicki, A. B. *et al. Am. J. hum. Genet.* **37**, 42–51 (1985).
9. Tuan, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 6937–6941 (1983).
10. Scharf, S., Horn, G. & Erlich, H. A. *Science* **223**, 1076–1078 (1986).
11. Saiki, R. K., Arnheim, N. & Erlich, H. A. *Biotechnology* **3**, 1008–1012 (1985).
12. Bugawan, T., Horn, G., Long, C. & Erlich, H. A. (Manuscript in preparation).
13. Horn, G. *et al.* (manuscript in preparation).
14. Chang, H., Moriuchi, T. & Silver, J. *Nature* **305**, 813–815 (1983).
15. Moriuchi, J., Moriuchi, T. & Silver, J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3420–3424 (1985).
16. Albert, E. D., Baur, M. P. & Mayr, W. R. *Histocompatibility Testing*, 4–8 (Springer, Berlin, 1984).
17. Reed, K. C. & Mann, D. A. *Nucleic Acids Res.* **13**, 7207–7221 (1985).

Common evolutionary origin of legume and non-legume plant haemoglobins

Jörg Landsmann, Elizabeth S. Dennis,
Thomas J. V. Higgins, Cyril A. Appleby,
Alexander A. Kortt* & W. James Peacock

Commonwealth Scientific and Industrial Research Organisation,
Division of Plant Industry, GPO Box 1600, Canberra,
ACT 2601, Australia

* Division of Protein Chemistry, CSIRO, Parkville,
Victoria 3052, Australia

The detection of haemoglobins in distantly-related non-legume plant families^{1,2} as well as in the legume families raises the question of whether the origin of the plant haemoglobins was single or multiple. We have isolated a haemoglobin gene from *Parasponia andersonii* (Ulmaceae) which shows more than 50% nucleotide sequence homology with the haemoglobin genes of legume plants and has three introns at identical positions to the leghaemoglobin introns^{3–5}. The *Parasponia* gene has homology to haemoglobin genes in another distantly-related nodulating plant, *Casuarina*. It also hybridizes at high stringency to sequences in a related, but non-nodulating genus, *Trema*. We conclude that the globin gene family may be widespread in modern plants, that plant haemoglobins may have a cryptic function in non-symbiotic tissue and that plant haemoglobins have evolved by vertical descent, probably from an ancestor common to modern plants and animals.

A number of plants outside the legume families are known to form nitrogen-fixing root nodules in symbiotic association with *Rhizobium*⁶ or actinomycetes⁷. Haemoglobin has been detected in nodules of several non-legume plants^{1,2} and characterized from *Parasponia* (Ulmaceae)^{8,9} and *Casuarina* (Casuarinaceae)¹⁰. Modern systems of plant classification show the known nodulating, haemoglobin-containing species to be widely distributed over several superorders of dicotyledons^{7,11,12} (Fig. 1). No continuous evolutionary pattern is evident.

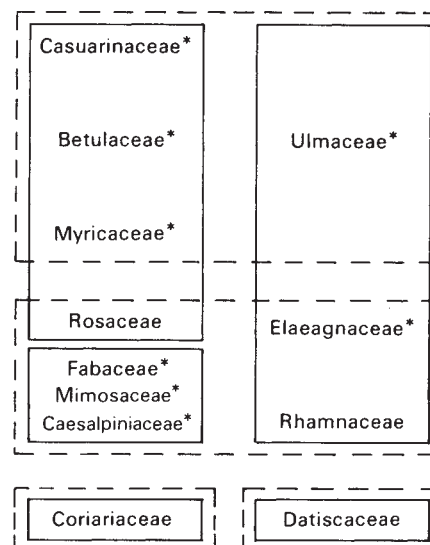


Fig. 1 Phylogenetic relationship of dicotyledonous plant families in which some species have been shown to form nitrogen-fixing nodules⁷. *Parasponia* is the only genus of the Ulmaceae which has been shown to nodulate, whereas in the legume families nearly all genera of the Fabaceae and Mimosaceae nodulate; the Caesalpinaceae are mostly non-nodulating. Families in which at least one species has been shown to contain nodule haemoglobin^{1,2} are marked with an asterisk. Solid line, superorders according to Dahlgren¹²; dotted line, according to Cronquist¹¹. Note that these superorders also contain ~150–250 (depending on the classification system) non-nodulating families not listed here. The N₂-fixing symbiosis in legumes is always with *Rhizobium* bacteria, *Parasponia* is the only known non-legume genus to be nodulated by *Rhizobium* strains, some of which also infect legumes⁶. Other non-legumes undergo symbiosis with *Frankia*, an actinomycete⁷.

Although there is 40% protein homology between soybean and *Parasponia* haemoglobin⁹, a soybean leghaemoglobin complementary DNA (cDNA) clone did not hybridize to *Parasponia* DNA or nodule RNA, even at reduced stringency. A 17mer and a 14mer oligonucleotide sequence deduced from the *P. andersonii* haemoglobin I protein sequence⁹ were made (see Fig. 4). Both oligonucleotides hybridized to an RNA 750 bases long in Northern blots of *Parasponia* nodules (Fig. 2a). The 17mer was used as primer for synthesis of radioactive cDNA, which was used as a probe; it also hybridized to a 750-base RNA in Northern blots (Fig. 2b) and identified haemoglobin-specific clones in a nodule cDNA library. Two positive clones were analysed; pL209 contained the entire coding region for *Parasponia* haemoglobin I whereas pL201 lacked the first 16 codons. Using pL209 as a probe, three independent genomic clones were isolated from a *Bam*HI-EMBL4 genomic library. All three clones contained one complete *Parasponia* haemoglobin gene on a 4.8 kilobase (kb) *Eco*RI end fragment within a 13-kb *Bam*HI fragment. The sequences of the genomic clones, pL305 and pL401, correspond to the cDNA clones pL209 and pL201 respectively (Fig. 3).

The coding region of the *Parasponia* haemoglobin gene spans 486 nucleotides which translate into 161 amino acids excluding the N-terminal Met (Fig. 4). Predictions from the DNA sequence showed discrepancies with the published protein sequence⁹. Resequencing of the relevant peptides confirmed an additional Phe at position 36, an additional Val at position 73 and a Cys at position 78 instead of the previously reported Thr. Four additional amino acids (Ser-Ser-Ser-Glu) were found at the N-terminus (Fig. 4). *Parasponia* haemoglobin has the widely conserved globin residues C2-Pro, CD1-Phe, proximal His (E7) and distal His (F8) and, as in other plant haemoglobins, an extended EF-region¹³. The protein coding DNA region of the *Parasponia* haemoglobin gene shows more than 50% homology

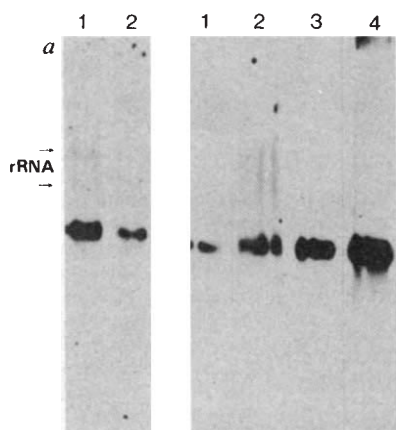


Fig. 2 Northern hybridization of *P. andersonii* nodule poly(A)⁺ RNA. *a*, Probed with: lane 1, the 17mer oligonucleotide; lane 2, the 14mer oligonucleotide. *b*, Probed with: lane 1, a mixture of 14mer and 17mer; lane 2, 17mer primed cDNA; lane 3, pL201; lane 4, pL209 (see Fig. 4 for sequences of the oligonucleotides). **Methods.** Aseptically-germinated seeds of *P. andersonii*, inoculated and nodulated with *Rhizobium* CP283, were transplanted and grown in a sand/vermiculite mixture and cultured using a nitrogen-deficient nutrient medium⁶. Root nodules were harvested from a number of plants and RNA was prepared from these according to standard procedures²². In Northern hybridizations 10 µg of poly(A)⁺ RNA was run on a formaldehyde agarose gel and hybridized in 50% formamide, 3.5×SSC, 0.75×Denhardt's, 15 mM NaPO₄ buffer pH 6.5 according to standard procedures²². Hybridizations with cDNA clones and 17mer-primed cDNA as probes were at 42 °C, but at 32 °C when oligonucleotides were used as probes. In oligonucleotide-primed cDNA synthesis, equimolar amounts of poly(A)⁺ RNA and 17mer were annealed for 30 min at 49 °C in 50 mM Tris/HCl pH 8.3, 50 mM KCl, 8 mM MgCl₂. The first strand of the cDNA was labelled with [α -³²P]dCTP using AMV-reverse transcriptase. Oligonucleotides were end labelled with [γ -³²P]ATP and T4 polynucleotide kinase.

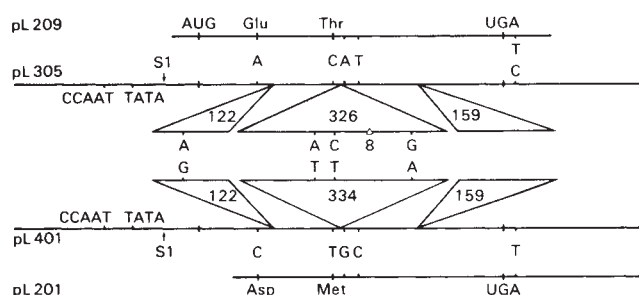


Fig. 3 Structure of cDNA clones pL209 and pL201 and genomic clones pL305 and pL401. The cDNA clones contain *Pst*I-*Pst*I inserts of 664 bp (pL201) and 613 bp (pL209). The genomic clones contain a 4.8-kb *Eco*RI fragment each. The DNA sequence of pL209 corresponds to pL305, the sequence of pL201 to pL401. Differing nucleotide positions and deduced amino-acid residues are indicated (see also Fig. 4). Translational start (ATG) and stop (TGA) codons and presumptive promoter signals for transcription (CCAAT and TATA) and S₁-mapped start of transcription are indicated. Intron 1 differs by 1 bp between the two genomic clones and intron 2 by 3 bp and an 8-bp deletion/insertion.

Methods. For cDNA cloning we followed the method of Gubler and Hoffman²³, using oligo(dT) as primer. After the final ethanol precipitation in 2 M NH₄-acetate the unfractionated double-stranded cDNA population was C-tailed with ³H-dCTP and terminal deoxynucleotidyl transferase, heated for 5 min at 65 °C and annealed for 4 h at 42° to G-tailed pUC8 containing about 12 G residues at *Pst*I ends (gift of W. Gerlach) according to standard procedures²². Transformation of *Escherichia coli* JM83²⁴ was according to Hanahan²⁵. Approximately 10⁴ white colonies per µg of input *Parasponia* poly(A)⁺ RNA were isolated on X-gal plates containing 100 µg ml⁻¹ ampicillin. Colonies were hybridized as described in Fig. 3 but with 5×SSC at 42°C with 17mer-primed cDNA as probe. Two positively hybridizing clones, pL201 and pL209, were analysed. The clones do not contain a poly(A) tail, probably due to the use of RNase H in cDNA cloning. A genomic library in λ EMBL4 was made from a 10–15-kb *Bam*HI-fraction of *P. andersonii* total DNA. This DNA was made from very young leaves taken from a number of plants. Plaques were hybridized with pL209 as probe. Three independent positive clones (out of 120,000) were identified. All contained a 13-kb *Bam*HI fragment containing a 4.8-kb haemoglobin-specific *Eco*RI-fragment. These fragments were subcloned into pUC19 from two of the genomic clones giving rise to pL305 and pL401.

with leghaemoglobin genes from soybean and kidney bean (Fig. 4). DNA homology is higher than the protein homology because amino-acid substitutions in various parts of the haemoglobin protein (for example, positions 18–34, 53–62, 71–78 and 91–100) have, in these regions, allowed mutations in the first codon positions to occur as frequently as in the second and third positions, thus altering the protein faster than its encoding DNA. However, the non-coding regions of the *Parasponia* haemoglobin gene have diverged from those of the leghaemoglobin genes. Only short sequences, which may be involved in regulatory functions, remain similar.

The two genomic *Parasponia* haemoglobin DNA sequences contain three introns (Figs 3 and 4) which are at the same positions as those in the leghaemoglobin genes. Introns 2 and 3 are inserted between codons; intron 1 splits a codon after its second base in *Parasponia*, as in the leghaemoglobins^{3–5} and in the vertebrate globins^{14,15}. The central intron is unique to the plant haemoglobins^{3–5,14,16}. The conservation of these intron positions in legumes and a non-legume plant indicates a common origin for plant haemoglobins from an ancestral functional gene unit having three introns. Gō suggested¹⁷, on the basis of protein domain analysis, that a third central intron would also have been present in the putative ancestor of animal globins.

In contrast with the haemoglobin multigene families found in legume plants³, Southern blot analyses of *Parasponia* genomic DNA indicate there is only one haemoglobin gene locus. A single band appears in digests with *Bam*HI (13 kb), *Eco*RI (7 kb), *Hind*III (9 kb) and *Bgl* II (8.5 kb) with pL209 as probe. However, two different haemoglobin clones were isolated from both cDNA and genomic DNA. The clones differ at four nucleo-

tide positions in the coding region, two of which led to an amino-acid change, and by a few bases in introns 1 and 2 (see Fig. 3). Intron 1 has 122 nucleotides in both genomic clones, intron 2 has 326 or 334 nucleotides and intron 3 is 159 nucleotides long. With the exception of a single base the clones are identical over at least 300 base pairs (bp) upstream and downstream in the non-translated sequences. These data suggest the existence of allelic heterogeneity at a single locus rather than different genes. The fact that *Parasponia* has not developed a multigene haemoglobin family as in the legumes, and that *Parasponia* haemoglobin has an unusually long polypeptide chain may reflect a relatively primitive haemoglobin gene condition¹⁸.

Legume and *Parasponia* haemoglobins both have around 15% protein homology with animal globins and introns 1 and 3 in plant haemoglobins are at positions identical to the introns in vertebrate globins. A monophyletic origin of all plant and animal haemoglobins and their vertical evolution seems probable. The invocation of horizontal gene transfer³ between animals and plants appears to be unnecessary but if it occurred it is likely that it predates the diversification of angiosperm plants¹⁹. A presumptive common ancestor of animal and plant haemoglobins could date back 1,500 million years, before the time of the radiation of animals and plants²⁰.

The *Parasponia* haemoglobin cDNA probe hybridizes to haemoglobin genes in the distantly-related *Casuarina*, which

Fig. 4 Comparison of the coding regions of the *P. andersonii* haemoglobin gene (sequence indicated is pL209/305, deviations seen in pL201/401 above the sequence) with those of kidney bean haemoglobin⁵. Spaces in these regions indicate identity of leghaemoglobin nucleotide sequence to *Parasponia*. The kidney bean introns (↔) are different in length, very low in homology and are not compared. The same applies to the 5' and 3' untranslated regions. Also indicated are the positions of oligonucleotides used in cDNA clone identification (14 mer mixture: CTC/TTACTTTTTCG; 17mer mixture: TTC/TAAGCTT-TACTTTGG). The derived (and corrected with respect to ref. 9) *P. andersonii* haemoglobin I amino-acid sequence is given in the bottom line. S₁ indicates mapped start of transcription. Presumptive promoter elements are underlined (----). **Methods.** DNA sequencing was according to Messing²⁴ following subcloning of appropriate fragments into M13 mp phages. Sites used for subcloning of cDNA clones were *Pst*I and *Sau*3A, and for genomic clones were *Eco*RI, *Hpa*I, *Bcl*I and *Sca*I. Synthetic oligonucleotides were made to facilitate sequencing of IVS-2 from both strands. The DNA sequences of pL201, pL209, pL305 and pL401 were determined at least twice from different M13 mp subclones.

We thank Ms G. Koci and Ms J. Norman for technical assistance, Dr M. J. Trinick for his supply of nodulated *Parasponia* seedlings, Dr B. Barlow for his advice in designing Fig. 1 and Drs A. and K. Marcker for fruitful discussions.

12. Dahlgren, R. M. T. *Bot. J. Linn. Soc.* **80**, 91-124 (1980).
13. Dickerson, R. E. & Geis, J. in *Hemoglobin: Structure, Function, Evolution and Pathology*, 68-69 (Benjamin/Cummings, California, 1983).
14. Blanchetot, A., Wilson, V., Wood, D. & Jeffreys, A. J. *Nature* **301**, 732-734 (1983).
15. Efstratiadis, A. *et al. Cell* **21**, 653-668 (1980).
16. Antoine, M. & Niessing, J. *Nature* **310**, 795-798 (1984).
17. Gö, M. *Nature* **291**, 90-92 (1981).
18. Zharkikh, A. A., Solovoyov, V. V. & Kolchanov, N. A. *J. molec. Evol.* **21**, 42-53 (1984).
19. Runnegar, B. *J. molec. Evol.* **21**, 33-41 (1984).
20. Dayhoff, M. O., Hunt, L. J., McCluchlin, P. J. & Jones, D. D. in *Atlas of Protein Sequence and Structure* (ed. Dayhoff, M. O.) 17-30 (National Biomedical Research Foundation, Washington, 1972).
21. Akkermans, A. D. L., Abdulkadir, S. & Trinick, M. J. *Pl. Soil* **49**, 711-715 (1978).
22. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
23. Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-269 (1983).
24. Messing, J. *Meth. Enzymol.* **101**, 20-78 (1983).
25. Hanahan D. *J. molec. Biol.* **166**, 557-580 (1983).

Study in indifference

Is the British public not interested in science books? Or is it being poorly served by the review media? John Manger of Oxford University Press on the trans-Atlantic contrasts in the reception accorded to a biography of Albert Einstein.

IN 1982 Oxford University Press published 'Subtle is the Lord' ... *The Science and the Life of Albert Einstein*, by Abraham Pais. It was judged to be, and remains, a superb piece of scholarship, but not without its problems. Why another Einstein book, asked the doubters, and so soon after the centenary celebration in 1979? We took the plunge anyway, in the belief that we had a truly original and saleable book on our hands. So it turned out to be. Over the next year or so over 30,000 hardback copies were bought world-wide, and subsequently a comparable number of copies of the paperback. The distribution of sales, and how the book was treated by the respective review media in Britain and the United States is illuminating.

Of the hardbacks, more than 25,000 were sold in the United States, the remainder in the rest of the world, about 2,500 of them in Britain. To put these figures in perspective, an average scholarly physics monograph would be expected to sell 1,000–1,500 copies in hardback world-wide, and more often than not would not be published in paperback. Most new novels sell far fewer than 30,000 copies in hardback.

On both sides of the Atlantic, science magazines and those publications largely or completely devoted to reviews did well by the book. In the United States it was reviewed by *The New York Review of Books*, *Science*, *Physics Today*, *Scientific American* and numerous other weeklies and journals; in Britain by *The Times Literary Supplement*, *The Times Higher Educational Supplement*, *Nature* and *New Scientist*.

In the United States, however, the book also attracted much wider attention — from *The New York Times*, *Los Angeles Times* and *Christian Science Monitor*, as well as from some 30 newspapers across the nation, ranging from the Youngstown (Ohio) *Vindicator* to the Phoenix (Arizona) *Republic*. The review in *The New York Times* was the lead, with picture, in the Sunday books section, perhaps the prime review position in the world. Booksellers, too, were enthusiastic in setting up window displays, while a science book club initially took several thousand copies and then came back for more.

None of the above publicity was particularly exceptional for an American trade publishing house. What was exceptional was the willingness of review editors in such a wide variety of newspapers and

magazines to devote space to such a book. It was, after all, not a simple layperson's guide to Einstein, but a scholarly, yet readable, account of the man and his not very simple science.

The book's reception by the "general" media in Britain was in stark contrast. Our approaches to the national weeklies and dailies for review coverage were, in general, met with bewilderment: "Science? We don't have anyone who can review science books"; or "Our readers aren't interested

stein, but he's not really very big over here". Who is big over here? In the week we would have expected reviews of 'Subtle is the Lord' the dailies and weeklies were obsessed (in terms of review space) by a volume of the diaries of Beatrice Webb. Many column inches, plus suitable period photographs, were devoted to the earnest, but ultimately trivial, meanderings of this fine Fabian. In the great scheme of things, it is difficult to believe that Einstein is not a more important figure to more people, even in Britain, than Beatrice Webb.

The literary editor would doubtless discount this as a somewhat tendentious conclusion, and retort that these kinds of judgements cannot and should not be made when it comes to deciding which books are reviewed in the national press. A literary editor must do what is best for

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Famous physicist and fine Fabian — Einstein (on a visit to the United States) and Beatrice Webb.

in science"; or "We have 20,000 books a year submitted to us for review, we can't cover them all, you know". The only national newspaper review appeared in *The Observer*, 17 column inches of it. Otherwise, except for BBC coverage in *Kaleidoscope* and on the World Service, the book was virtually ignored. There were no book club sales (there is no scientific book club in Britain).

Two questions arise from this experience. Is the British public not interested in science books? Or is it being poorly served by the review media? Bearing on the first of these is the existence in the United States of the well-known and popular "Physics for Poets" course. Students from most disciplines are encouraged to take at least one science option, and physics "without the sums" is widely taught. Stimulating textbooks are produced in large quantities by college publishers and the end result is a sizeable population of students who have a knowledge of space, time, the nature of the Universe, the atom, crystal structure and, yes, even relativity. More importantly, individuals are not encouraged to think it is somehow smart to be scientifically illiterate.

They will certainly also have an awareness of the importance of a character such as Einstein. One of the more amusing responses to our attempts to attract attention to Pais's book was "Well you'd expect the Americans to be interested in Ein-

circulation — quite right. But is it a valid assumption that 99 per cent of the readers of the daily and weekly papers have no interest in science for 99 per cent of the time? Because that is the legitimate conclusion to be drawn from the amount of review space given to science books. It should not be forgotten that not only have the social implications of science never been greater, but that there are hundreds of thousands of practising scientists in Britain some of whom can even read newspapers.

Per head of population, the biography of Einstein sold twice as well in the United States as it did in Britain. But it is nonetheless almost certainly true that there is a greater interest in science in Britain than the media lead themselves to believe. The worrying corollary of their indifference (or plain ignorance) is that publishers will tend to take fewer risks, look for fewer new authors and develop fewer new talents if they feel that science books will be "published without trace". The example of Pais's book demonstrates that science books do not have to be trivial to be saleable in global terms. The pity of it is that the review media in Britain remain stubbornly resistant to educating themselves or their readers. □

John Manger is Editorial Director responsible for science and medicine at Clarendon Press, an imprint of Oxford University Press, Walton Street, Oxford OX2 6DP, UK.

One-eyed darwinism

Philip Kitcher

The Blind Watchmaker. By Richard Dawkins. Longman: 1986. Pp. 332. £12.95.
To be published in the United States on 24 November by W.W. Norton, \$18.95.

To Richard Dawkins it is always *the* problem. How are we to account for the complex design in nature, the intricacy of fit between organism and environment that once moved William Paley to propose the analogy of finding a watch on the beach, and thereupon build an argument for the existence of a Creator? Dawkins admires Paley's appreciation of a great mystery, even though he has no sympathy for Paley's theological solution. Thanks to Darwin, we have no need of that hypothesis. The preordained unfolding of the Creator's plan is replaced by the action of natural selection, the blind watchmaker of Dawkins's title.

Dawkins believes, with considerable justice, that the theory of evolution by natural selection is widely misunderstood, and part of his project is to explain that theory in a manner that will demonstrate its power to solve Paley's puzzle. *The Blind Watchmaker* contains passages of exposition that deserve sustained applause. The third chapter begins by distinguishing clearly between single-step selection and cumulative selection, and proceeds through some ingenious and illuminating computer simulations which make plain how cumulative selection can give rise to striking and unanticipated modifications of an original form. It is tempting to suggest that Dawkins's program should be made widely available, to schoolchildren and to eminent astrophysicists, so that the rumbling complaint that compares the effects of selection to a tornado sweeping through a junkyard may finally be laid to rest.

No less brilliant is the discussion at the end of Chapter 4, in which Dawkins the sensitive naturalist replaces Dawkins the clever computer hacker. In 15 pages, he reviews the standard examples of convergent evolution (and some lesser known cases) in prose that makes them extraordinarily vivid. Chapter 5 provides a clear and entertaining account of the elements of contemporary genetics, and Chapter 6 an outline of Cairns-Smith's hypothesis about the origins of life and a useful response to those who claim that the occurrence of life on our planet is so improbable

that it signals a miracle. The next two chapters reveal the ways in which natural selection can work constructively to yield "a building up of complexity that has more in common with addition than with subtraction" (p.169). Particularly admirable is Dawkins's lucid presentation of Lande's complex ideas about runaway sexual selection, a discussion in which Dawkins

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

was assisted by Alan Grafen.

The six central chapters of *The Blind Watchmaker* thus constitute a compelling exposition of the main themes of contemporary neo-darwinism, indeed the most compelling presentation of those themes that I know. The stylistic talents that won a broad readership for *The Selfish Gene* are fully in evidence here. But Dawkins aims not only to present what is uncontroversial amongst evolutionary biologists but also to take a stand on matters that continue to spark professional debate. In the end, *The Blind Watchmaker* tries to articulate a particular version of neo-darwinism, and in this part of his enterprise Dawkins is far less successful.

The last three chapters take up three main areas of disagreement in contemporary evolutionary theory. Dawkins considers the proposal that the theory of punctuated equilibria provides an expansion of neo-darwinism, the continuing controversies in systematics, and an assortment of positions that emphasize alternatives to selection as agents of evolutionary change. In this part of the book the real issues are never joined. Dawkins sees clearly that some of the most strident criticisms of neo-darwinism caricature the position that they advertise themselves as attacking. But his own defence is to identify untenable extremes and to portray the version of neo-darwinism that he favours as occupying the sensible middle ground.

I will concentrate on Dawkins's attempt to deflate the punctuationalist challenge to neo-darwinism. His strategy here is to analyse the three main claims advanced by the punctuationalists, showing that, under analysis, Gould, Eldredge, Stanley and others do not stray very far from neo-darwinian orthodoxy. The success of this strategy depends on how far the analysis proceeds. Eldredge and Gould made a proposal about the geometry of phylogenetic trees: evolution consists of long periods of stasis punctuated by short bursts of speciation. Dawkins fastens on the idea that speciation occurs "suddenly", and argues that the rates envisaged are consonant with those admitted by gradualists. But this is to bypass the main point of difference. Punctuationalists contend that evolutionary changes are concentrated around episodes of speciation, and they are (to say the least) sceptical about the importance of either morphological change or speciation within unbroken lineages.

In discussing punctuationalist claims about the mechanisms underlying stasis and speciation,

Dawkins also finds only what is orthodox. The idea that small molecular substitutions may have large phenotypic effects is, he claims, a part of mainstream neo-darwinism. Thus, the punctuationalist emphasis on the importance of modified developmental pathways in speciation can be assimilated into the reigning orthodoxy. The issue here is closely related to that in the chapter where Dawkins considers the complaint that neo-darwinians overemphasize the importance of natural selection in evolution and fail to appreciate the role of developmental constraints.

There, Dawkins concludes by claiming that "all serious Darwinians" would agree that "we can't afford to ignore the con-

straints on evolution that embryology imposes" (p.311). But although, like Dawkins, most neo-darwinians are now prepared to declare that the power of selection to shape a phenotypic trait is limited by developmental factors, the critics of orthodox neo-darwinism argue that in practice this declaration is forgotten and that explanations of phenotypic traits effectively suppose that natural selection is omnipotent. The debate should centre not on the extreme "Mutationist" who figures in Dawkins's pages, but on the relative importance of various agents of evolutionary change.

Dawkins's failure to identify the real question in this case is reflected in a failure to deal adequately with the punctuationists' account of stasis, and he is led to canvass extraordinary possibilities in interpreting their views. Allegedly, punctuationists believe that species "actively resist change" (p.247), and Dawkins has little trouble in rebutting one obvious reading of this phrase. Charity easily finds alternative interpretations. If there are indeed strong developmental constraints, then throughout most of the history of a species there may be very few options to exercise selection. Only when the environment changes radically, disrupting the mechanisms that typically buffer ontogenesis, will hitherto cryptic variation become apparent and will there be a possibility for evolutionary experimentation. Whether or not scenarios of this kind prevail in the history of life, the claim that they do represents an important challenge to the vision of stately unfolding that permeates orthodox neo-darwinism.

Dawkins's discussion of the third plank in the punctuationists' platform, the thesis that many macroevolutionary trends are the result of processes of species selection, is extremely brief, and the brevity is significant. Instead of presenting his reasons for thinking that species selection is not very important in evolution, Dawkins refers the reader to his previous book, *The Extended Phenotype*. This was a wise decision that would have been advisable at many other places in the last three chapters of *The Blind Watchmaker*, for *The Extended Phenotype* treats the leading controversies in a far more satisfying way than does the new book. However Dawkins does add to his reference another argument that is in tune with the main theme of *The Blind Watchmaker*. Species selection, he contends, cannot account for complex adaptations.

At this point, the sources of the successes and the shortcomings of *The Blind Watchmaker* click into place. Punctuationists will want to combat the argument, alluded to but not presented, for the thesis that species selection is unimportant in evolution, but they may be quite unworried by the charge that species selection cannot account for complex adaptations.

Part of their case is that the prevalence of complex adaptations in the history of life has been exaggerated by orthodox neo-darwinians. *The Blind Watchmaker*, which makes the explanation of complex adaptations central to the neo-darwinian enterprise, sets up from the beginning a framework in which the character of the interesting debates within contemporary evolutionary theory cannot be made clear.

The book starts with an attempt to present Paley's puzzle. Dawkins would like his readers to feel the grip of the problem of explaining organic complexity, and he would like to present that problem in a way that does not prefigure the darwinian solution. But how exactly is the notion of a well-designed, complex object to be characterized? Dawkins displays great ingenuity in his effort to avoid the trivializing answer that explicitly defines a complex adaptation as the product of a history of selection. For those who view neo-darwinism as flawed by its overemphasis on adaptation, his struggles are likely to seem hopeless. They will suggest that the problem is to define a class which will turn out to include just those traits whose presence is to be understood in terms of a history of individual selection without making use of ideas that would have been unavailable to Paley (or any other pre-darwinian impressed with the apparent design of organisms). But, claim the anti-adaptationists, cases of apparent design are a motley. Some, but by no means all of them, are to be understood in terms of the action of individual selection. Only by introducing the concept of selection explicitly into our characterization of the class of "well-designed complex objects" can we manage to present Paley's problem in a way that allows for Dawkins's preferred darwinian solution. Hence, Dawkins's emphasis on the evolutionary explanation of "complexity" disguises an unargued commitment to the view that the problem of adaptation is central to evolutionary theory, and a corresponding slighting of important alternatives. For Richard Dawkins — but not for all sophisticated evolutionists and, arguably, not for Darwin — it is always the problem.

For all my reservations, I am confident that *The Blind Watchmaker* will do much good. At its heart is a remarkable account of some of the most important ideas in the history of science. Dawkins should clear up many more misapprehensions than he fosters, and his readers should gain a sense of the excitement of contemporary evolutionary biology. If they are moved to read further, then their initial acquisition of prejudices is likely to be corrected. The details, we may hope, will come later. □

Philip Kitcher is a Professor in the Department of Philosophy, University of California, San Diego, La Jolla, California 92093, USA, and author of Vaulting Ambition: Sociobiology and the Quest for Human Nature (MIT Press, 1985).



Mammal Evolution: an illustrated guide

R.J.G. Savage & M.R. Long

Unique colour restorations and other illustrative material combine with Professor Savage's authoritative text to provide a wide-ranging study of these diverse animals.

256pp, 120 colour, 120 b/w illustrations, 300 line drawings.
Hbk. 0 565 00942 7 **£17.50**

The Ancestry of The Vertebrates

R.P.S. Jefferies

The problem of the origin of the vertebrates is solved by deriving them from calcichordates, fossils found in marine rocks of Cambrian to Devonian age.

300pp, 300 b/w illustrations.
Hbk. 0 565 00887 0 **£50.00**

A Manual of Forensic Entomology

K.G.V. Smith

An important new book in this neglected area of entomology. Will be of use to forensic scientists, ecologists and archaeologists.

256pp, 402 figs.
Hbk. 0 565 00990 7 **£17.50**

Jobling's Anatomical Drawings of Biting Flies

Ed. D.J. Lewis

356 of Boris Jobling's drawings are published in this volume for the first time, the originals being held by the Wellcome Museum of Medical Science. Will be of particular interest to medical entomologists.

96pp, 96 plates.
Hardback. 0 565 01004 2 **£9.95**

Sales Department, Natural History Museum Publications, Cromwell Road, London SW7 5BD

What the scientist has to say

Arthur Kornberg

Storm Over Biology: Essays on Science, Sentiment, and Public Policy. By Bernard D. Davis. Prometheus, Buffalo, New York: 1986. Pp.324. \$22.95.

SCIENTISTS make poor politicians. They have been self-selected for their interest in things rather than people, and they are generally neither literary nor articulate. They form societies to communicate with one another about their arcane findings but, with some exceptions, they are distinguished from most other professional groups by their social isolation and political innocence.

Were the responsibility of scientists limited to progress in such areas as chess or hieroglyphics, their ineptitude in worldly matters might be tolerated with whimsical affection. But the affairs of government and business, our everyday lives and future depend crucially on science and its technological applications. Must we then rely solely on salesmen, actors and lawyers to make the difficult decisions that require informed judgement about scientific questions? What are we to do?

To compound the problem, the media seem to be entranced by the pronouncements of those who lack scientific knowledge and of a few scientists who have ideological biases and crave publicity. The brushfire controversies over drugs, toxic wastes and genetic engineering have coalesced into wider conflagrations of hostility to all technological, scientific and intellectual effort. Legitimate concern about biological warfare has turned into all-out warfare against biology.

Fortunately there are a few informed and responsible scientists who speak out and write clearly on issues involving biology and society. One such is Bernard D. Davis. In *Storm Over Biology* he has collected together 44 essays and shorter pieces on "science, sentiment and public policy". All but one of these have appeared in print in the past ten years, many in obscure places. Trained as a physician, now a scientist with a distinguished record of administration and of research in biochemistry and microbiology, Davis has intimate knowledge of the difficult issues that he has chosen to confront.

Among the matters that get attention are sociobiology, affirmative action in medical schools and genetic engineering. The common theme underlying all of these topics is the need to build on reality and protect science from politicization by the left or the right. In particular, on the perennial debate over biological versus

social determinism ("nature versus nurture"), Davis makes cogent arguments for rigorous studies of human genetics and behaviour against wailings by political ideologues about the potential abuses of new knowledge.

In his discussions of the applications of recombinant DNA technology, Davis speaks from broad experience in microbial physiology, genetics and evolution. Recognizing the legitimacy of concern about the hazards of genetic engineering, he offers convincing arguments for the improbability of the scenarios that have most alarmed the public. He further argues for caution in genetic intervention in man, and for recognizing the limits to the possibilities for such intervention. Against specious attacks by Jeremy Rifkin, recently abetted by the editorial board of *The New York Times*, he takes a bold stand for both truth and reason.

A particularly moving, previously unpublished piece deals with the history of the response to his pleas for balance between affirmative action and maintaining academic standards in medical schools. Davis — a child of immigrants, a witness to anti-Semitism, a perennial

fighter for liberal causes and the first department chairman in the history of Harvard University to appoint a black to tenure — cannot be accused of insensitivity to the need for social justice. Yet his editorial in the *New England Journal of Medicine*, arguing for reasonable stretching in increasing the opportunities for minority candidates, was presented by the news media as a racist attack, and he was vilified by many colleagues at Harvard Medical School in the hysterical atmosphere that swept through academic faculties during the late 1960s and early 1970s.

I sometimes wished for the dashes of humour that lace Davis's lectures, but the issues in these essays are really serious business. In his foreword, Edward Shils suggests that scientists generally underestimate the depth of the current anti-science movement, and he makes it clear why this important set of statements by Davis will reward the attention of the general public as well as scientists. □

Arthur Kornberg is a Professor in the Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, USA.

Sigmund meets Charles

John R. Durant

The Problem of Altruism: Freudian-Darwinian Solutions. By C. R. Badcock. Basil Blackwell: 1986. Pp. 206. £14.95, \$24.95.

BOTH natural and social scientists regard Freud's psychoanalytic theories as a somewhat baroque construction of over-elaborate and largely untestable speculations. *The Problem of Altruism* is dedicated to the task of restoring them to their rightful place in the scientific sun. Christopher Badcock is a social scientist who regards conventional social theories of all sorts as hopelessly inadequate to their task. These theories, he argues, being committed to various forms of "holism" and "cultural determinism", are all involved in the futile attempt to explain one social phenomenon in terms of another. What is needed instead, he suggests, is a reconstruction of sociological theory on more reductionist and individualist foundations, and he attempts to lay these foundations by combining psychoanalysis with darwinian social theory, or human sociobiology.

At first sight, this seems a most unpromising strategy. Not only does human sociobiology have a reputation of its own for over-elaboration and untestability, but also it appears to be quite distinct in its

approach to the explanation of behaviour. Where freudian theory deals with psychological entities such as the conscious, the pre-conscious and the unconscious minds, and with psychological processes such as projection, regression and repression, darwinian social theory deals with biological entities such as genotypes and phenotypes, and with biological processes such as kin selection and reciprocal altruism. Where is the room for marriage between these two?

Badcock's central argument is that the place where psychoanalysis and sociobiology meet is in the field of "dynamic psychology". For example, the American sociobiologist Robert Trivers has analysed the relationship between parents and their offspring in terms of psychological conflict; and he has debated the possible adaptive advantages of the unconscious mind. In each case, a recent sociobiological analysis appears to mirror an older, psychoanalytic view. Modern darwinists, Badcock suggests, are busy rediscovering and vindicating Freud.

Badcock builds his case for the rehabilitation of Freud around the central theoretical problem of sociobiology, namely altruism. Sociobiologists admit only two major mechanisms for the evolution of behaviour which raises the reproductive fitness of another individual at the expense of that of the actor: reciprocity, and kin selection. Reciprocity, he suggests, may have become an important element in human social behaviour with the development of cooperative big-game hunting; and it may have led to the rapid evolution

of complex psychological mechanisms enabling individuals better to practise and detect cheating, including elaborate moral codes and the unconscious. Similarly, kin selection may have promoted the evolution of a basic psychological mechanism of identification; and once in place, this may have permitted the development of anything and everything from masochism to heroic self-sacrifice on the field of battle.

The strength of this freudian-darwinian approach to human behaviour lies in its recognition of the need for a cogent psychological theory to place between the abstract principles of biology, on the one hand, and the concrete realities of human society, on the other. Its weakness, however, lies in the sloppiness of the theoretical constructs that lie at the heart of darwinized psychoanalysis. Take, for example, the central notion of identification. Supposedly, our ancestors evolved a basic ability to identify with close kin in order that they should aid them in circumstances that would promote their own inclusive fitness. Well and good; but for such an ability to evolve, it is necessary that it should be highly discriminating; for much (indeed, most) helping behaviour is detrimental to inclusive fitness and could never evolve. What we need to know, then, is what kind of identification mechanism is being proposed. This we never learn. Instead, we are presented with an allegedly evolved genetic-psychological trait which is so indiscriminate that it leads people to behave altruistically towards virtually anyone and everyone. Under what circumstances could such a dangerous trait as this possibly evolve?

Psychoanalytic theory has always suffered from the problem that its central claims are insufficiently specified to permit the formulation of scientifically testable predictions. Recently, a certain genre of sociobiological literature has rediscovered the superficial appeal of equally vaguely-formulated hypotheses about the roots of human conduct. Having set out to combine the two enterprises, it is perhaps not surprising that Badcock has fallen victim to the trap of explaining both too little and too much. This is a pity, for the ideal of a powerful and a predictive evolutionary psychology is a worthy one. We know that Darwin's basic insights can be applied rigorously to behaviour in general. Perhaps at least some of Freud's ideas could be re-cast in such a way that they, too, could be applied rigorously to human conduct; but this has yet to be shown. In the meantime it has to be said that no genuine "freudian-darwinian solutions" are to hand. □

John R. Durant is Staff Tutor in Biological Sciences in the Department for External Studies, University of Oxford, Rewley House, 1 Wellington Square, Oxford OX1 2JA, UK.

Scrotum hokum

Walter Gratzer

The Monkey Gland Affair. By David Hamilton. Chatto & Windus, London: 1986. Pp.155. £11.95.

ACCORDING to Dr Johnson, a life protracted is protracted woe. In 1953 the good Pope Pius nevertheless journeyed to Switzerland to receive injections of testicular tissue at the hands of the egregious Dr Niehans. His bid for rejuvenation (though presumably not virility, by which it was generally measured) must have been accounted successful, for shortly afterwards the Mephistopheles in this Faustian contract was elevated to membership of the Pontifical Academy. This occurred two years after Serge Voronoff, in his day the most celebrated and successful exponent of testicle (or, as the euphemism of the time had it, monkey gland) implants, had died, discredited, despised and mocked by the society that had lionized him for two decades. Science does not eradicate human credulity, it merely shifts its focus.

The search for the secret of eternal youth has attracted not only crackpots and mountebanks, but also serious and sometimes even distinguished scientists, whose reason could be selectively suspended. The most illustrious of these was Metchnikoff, who was convinced that aging was caused by bacteria in the gut, which destroyed connective tissue. These, he held, could be driven out by abundant ingestion of yoghurt. Examples nearer to our own time include the view that antioxidants can arrest the processes of decline through the elimination of free radicals, and that lead clothing to exclude cosmic radiation will encompass the same end. None of these theories ever achieved much of a popular following; yet in the 1920s one

fancies, there must have been queues of rich, aging voluptuaries in the ante-rooms of "Doc" Brinkley's clinic in Kansas and at the Vienna "Vivarium", their trousers draped over their arms, braces dangling, impatiently awaiting their goat testis implants or unilateral vasectomies — which the learned Professor Steinach was able to assure them would result in the retention of sperm in the testis and thus suffusion of the system with youth and vitality. (This is curiously reminiscent of the principle upheld by the yogis, that the emission of sperm, which they equate with manly virtue, enfeebles the body, so that one must learn to aspirate it back after coition. Many Victorian moralists subscribed to the same opinions, but prescribed abstinence as a corrective.) Steinach treated women by the even less appealing method of X-irradiation of the ovaries.

David Hamilton, who is both a surgeon and a historian of medicine, has written a scholarly, absorbing and vivid account of the history of testicular implantation, which thrived in the first half of this century. For Hamilton, the most interesting figure amongst the *galère* of improbable characters that throng his pages is Voronoff, because he was sincere in his efforts, deluded rather than a charlatan, and within limits rational in his beliefs.

These limits were set by an unfathomed though common ignorance, verging on obscurantism, about the techniques of experimental design and evaluation. The same attitudes of mind are still with us, and they animate homeopathy, psychoanalysis and a range of other peripheral disciplines. The reasoning behind implantation therapy was not new, and amounted to little more than that the testis was the seat of fertility, therefore of youth. However the discovery of endocrine secretions encouraged the plausible hope that foreign testicular tissue might be the source of some benign active princ-



The Kansas quack — a feature about "Doc" Brinkley in the New York Evening Journal. Brinkley is shown holding the first "goat gland baby", born to one of the recipients of a goat testis implant.

ple. Voronoff became persuaded that a successful xenograft must in general come from a closely related species, not for example from ruminants, and many monkeys were emasculated in the interests of this theory. In the United States prison doctors joined in the game, and were able to go one better: Dr Stanley of San Quentin gaol implanted 30 of his charges with testicles from executed felons and later reported highly satisfactory results from 643 assorted gland transplant operations. All these grafts must of course have been rapidly rejected.

Worse yet was Brinkley, the Kansas quack — unencumbered by any genuine medical qualifications. He chose the goat (an odourless variety), as a species traditionally associated with sexual exuberance, to turn the prematurely enfeebled into satyrs. He and his wife introduced what amounted to a conveyor-belt system of surgery. As a decorative touch, he even injected mercurochrome into the severed end of the vas deferens, to impart a rich colour to the urine and impress the patient. Brinkley owned the hotels that served as convalescent centres, and he set up a radio station to advertise his operation (which was offered in standard and *de luxe* versions), as well as a range of pharmaceutical products. He became extremely rich and popular, and fought a close-run election for State Governor before his luck ran out.

Meanwhile in Paris, Voronoff had married into money on a grand scale and had become a social lion. His second marriage was to a surmised daughter of King Carol of Roumania and the infamous Madame Lupescu. Hamilton makes the point that in the aftermath of the Great War, in which the best had perished and the breeding stock of the upper classes been seriously depleted, and when indeed eugenic theories were widely touted, European society was peculiarly receptive to promises of increased virility and fertility. Voronoff, who was a tireless and prodigiously talented self-publicist, took full advantage of these preoccupations. He also extended his activities to the breeding of sheep, in which he saw great pecuniary promise. The results of Lamarckian (or Lysenkoist) experiments, conducted in Algeria, on the effects of testicular grafts on the performance of rams and the wool-yield of their offspring, were proclaimed a triumph; the British Thoroughbred Breeders Association took fright and announced that grafted racehorses would be struck from the stud-book.

Then Voronoff fell like Lucifer. He was brought down largely through the resolute scepticism of a number of French and English vets, who knew an uncontrolled experiment when they saw one — a circumstance shaming to the medical establishment, which had vacillated feebly throughout Voronoff's years of ascenden-

cy. The most tenacious and clear-sighted was Henri Velu, who ran a sheep-breeding station in Morocco, and in the end he outfaced Voronoff and his many powerful supporters. Moreover testosterone had shortly before been isolated by (as Hamilton calls him) a Dr Funk. This was none other than Casimir Funk, the man who gave us the word vitamin and was said to have enraged the Mussolini regime when, learning that Italian prisoners taken in the Abyssinian War were being castrated by their captors, he attempted to negotiate a consignment of the resulting material. Injections of testosterone were found ineffective in the treatment of sexual decline. So began Voronoff's years of eclipse.

David Hamilton has penetrated the shell of Voronoff's formidable public persona, and come to like the man within, self-seeking *poseur* though he quite evidently was. *Tout comprendre*, as they say, *est tout pardonner*. The book is dedicated to Serge-Samuel Voronoff. □

Walter Gratzer is in the Medical Research Council Cell Biophysics Unit, Kings College London (KQC), 26–29 Drury Lane, London WC2B 5RL, UK.

Medical matters

Lewis Thomas

The Oxford Companion to Medicine, Vols I and II. Edited by John Walton, Paul B. Beeson and Ronald Bodley Scott. Oxford University Press: 1986. Pp.1,524. £55, \$95.

I PICKED up the first of these two volumes at 11 o'clock one evening, and could not lay it down until 11:05. This sounds like the start of an uncordial review, but it means only to say quickly that this new encyclopaedia is not made for steady late-night reading. But for leafing through at random, in spare moments, with nothing particular in mind, *The Oxford Companion to Medicine* contains some unexpected pleasures. All in all, it is an earnest but good-natured piece of work, 1,500 or so pages of what seem rather randomly selected topics, some of them just a few lines of definition, others lengthy essays on one or another aspect of medicine, health, disease, medical practice, and quite a lot on the organization and economics of medicine.

Among the surprises, for me anyway, is a long essay on Doctors as Patients ("doctors have a healthy skepticism about the efficacy of the art they practice"), Doctors as Athletes, Doctors as Missionaries, Doctors in Other Walks of Life (the arts, the law, philosophy, the Church, crime).

Although the editors state in their preface that they hope for a readership made up of intelligent laymen as well as physi-

cians, nurses and other health care professionals, including students, my guess is that the volumes will be most appealing to people who simply like browsing, whatever their background. If, as the editors hope, there is "a plastic surgeon wishing to learn more about the role and responsibilities of a psychiatrist", he will find plenty to read, eleven pages on Psychiatry and another ten on Psychology in Relation to Medicine. The psychiatrist, for his part, will find five pages on Plastic and Maxillo-facial Surgery, directly following a short paragraph explaining Plaster of Paris. And any practitioner or layman wanting to know more about Death, Dying and the Hospice Movement, or about Government and Medicine in the UK (and also in the USA), or about Community Medicine, or Dentistry, or Cardiothoracic Surgery, or Medical Museums, or the Medical Services of the Royal Navy, or Women in Medicine, will find pages of essays on these and other broad topics. But, along the way, the casual reader will run into the Opening Snap ("associated with mitral valve opening at the beginning of diastole"), Suffocation ("the interruption of breathing by deprivation of air"), Weakness ("lack of strength, feebleness, ill health"), Neurotoxin ("any agent which is toxic to nervous tissue") and a lot of other items not needing even the one or two lines given them.

Whatever the intended readership, one thing *The Oxford Companion to Medicine* is not: it is not, despite the title, any kind of guide for do-it-yourself-at-home medicine, and the lay person who shells out £55 with this in mind will be disappointed. There is, to my own disappointment, not even much material to enlighten the layman on what to stay away from, in the light of those endless shelves of paperbacks in every airport lounge providing promises by immunity-stimulating diets, magical exercises, trace metals, vitamins, rejuvenating lotions, even books on how to think straight. To be sure, if you go looking you can find some of this, under Fringe Medicine, Cults and Quackery. In the absence of an index (which these two volumes could surely use) you have to be lucky and browsing through the F's to find it, and even so this section deals mostly with the health frauds of the nineteenth century and not enough with the more outrageous follies of this one.

In short, this is a curious, rather unpredictable, almost quirky encyclopaedia. Some will find it useful in spots, interesting for some of the essays. I'm glad to possess it, but I'm afraid that long before my pages became thumbworn my thumbs would be pawgorn. □

Lewis Thomas is Professor Emeritus at the Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA, and University Professor at the State University of New York, Stony Brook.

Evolution in the twentieth century

Joe Felsenstein

Sewall Wright and Evolutionary Biology. By William B. Provine. *University of Chicago Press*: 1986. Pp. 545. \$30, £25.50.
Evolution: Selected Papers by Sewall Wright. Edited and introduced by William B. Provine. *University of Chicago Press*: 1986. Pp. 649. Hbk \$70, £59.50; pbk \$25, £21.25.

EVOLUTIONARY theory in the 1960s and 1970s was dominated by optimization, by the search for the quantity evolution would maximize. The mean fitness of a population was not precisely what natural selection would maximize in a Mendelian genetic system, so perhaps if we did enough theory we could find out what was. The picture hanging on our wall was of R.A. Fisher, who had emphasized the importance of natural selection in large populations. The contributions of Sewall Wright were acknowledged by all, but were considered to be of secondary importance.

Theoretical population genetics then saw itself as a branch of applied mathematics, and in need of more mathematical rigour. The only data to which we paid attention were gene frequencies of electrophoretic polymorphisms. The kinds of data which morphologists, palaeontologists and most systematists collected were ignored.

By the end of the 1970s many of these evolutionists working above the population level saw the neo-darwinian synthesis as irrelevant to their needs. Punctuated equilibrium and even neo-lamarckian alternatives gained adherents. The resulting controversies over punctuated patterns of evolution and species selection mechanisms have revived interest in complex fitness surfaces and small populations. The evolutionary views of George Gaylord Simpson and Ernst Mayr, both influenced by Wright, have been important in these debates. All of which has led to a remarkable resurgence of the reputation of Sewall Wright.

It is unusual for the details of a theoretician's work to be so relevant 70 years after his views were formed. It is even more unusual for a major account of his work and its influence to appear in his own lifetime. Provine's thorough and thoroughly admirable examination of Wright's life and influence, which is accompanied by a very useful collection of Wright's papers on evolution, is the best we have for any recent figure in evolutionary biology.

Sewall Wright was born in 1889, and grew up in Galesburg, Illinois, one of a generation of talented and energetic intel-

lectuals from the prairies of Illinois and Wisconsin. One is reminded of people such as the poet Carl Sandburg and the photographer Edward Steichen; in fact, Sandburg married Steichen's sister, and Wright's father Philip was Sandburg's most influential teacher and first publisher (Sewall set the type and ran the press). Provine gives careful attention to these early years, and to the intellectual atmosphere in which Wright trained; his chapters on Harvard's Bussey Institution and on evolutionary theory after Darwin are worth the price of admission by themselves. Here, I was startled to discover how closely the present controversies over the relative importance of species selection and individual selection were paralleled by controversies at the turn of the century.

Provine also goes into detail — more



Early days — Sewall Wright in 1922.

than seems readable — on the physiological genetics which was Wright's main concern until his retirement. As a student learning two fascinating stories, on mammalian coat colours and mathematical genetics, I did not appreciate how closely they were bound together. One is the story of explaining a seemingly arbitrary array of colours by the concentrations and properties of enzymes. Part of its fascination is evolutionary: the appearance of the same genetic variants in different mammalian species (Himalayan in rabbits, Chinchilla in mice and Siamese in cats, for example). The other extends the genetic system to the population level, making predictions from the necessary regularities of the system of inheritance.

Only later did I come to realize that the same person, Sewall Wright, was central to both stories, and that coat-colour genetics was critical to the formation of his views on evolution. A list of Wright's PhD students (pp. 189–190) discloses few who concerned themselves with population genetics, but makes his influence on

American mammalian genetics palpable. He wrote major reviews of physiological genetics, and influenced Beadle and Tatum's famous work on *Neurospora* which led to the one gene — one enzyme hypothesis. Although Provine deals chiefly with Wright's evolutionary theories, he recognizes the importance of the work on gene action in coat colour in informing Wright's view of the role of gene interaction in the inheritance of quantitative characters.

After describing Wright's years at Harvard (1912–1915), the US Department of Agriculture (1915–1925) and the University of Chicago (1925–1955), Provine devotes most of the rest of the book to Wright's views on evolution and their influence. There are some surprises. For example, in the course of an extended consideration of Wright's "shifting balance" theory of evolution Provine argues that Wright changed his views substantially in the 1940s. Wright himself denies this, but Provine's evidence, largely drawn from Wright's papers, seems incontrovertible. Before 1940, he points out, systematists believed that differences between species were of little or no adaptive significance. Wright's theory, originally developed with animal breeds in mind, could be applied to natural populations by supposing that random differences arose between species by genetic drift, and that selection between species utilized these to produce a net adaptive response. Provine believes that evolutionary biologists during the 1920s and 1930s did not understand the role of between-group selection in Wright's theory, that they understood him to be saying that genetic drift was the primary agent of evolution.

After 1940, as systematists and population geneticists reinterpreted species differences as being largely adaptive, Wright modified his theory. He brought it closer to the original animal-breeding model by having the random differences arise between local populations rather than between species. Provine finds evidence that this unadvertised change of views was a source of much confusion, playing a role in the famous post-war disputes with Fisher and Ford.

The controversies with Fisher are given careful attention. Provine takes pains to point out the reasons why these two came to disagree so completely about the evolutionary forces operating in nature, despite their agreement on most of the mathematical details. As long as we lack a comprehensive picture of the patterns of gene interaction and population structure in a wide range of species, the views of Fisher and Wright will form opposite poles of microevolutionary theory. One pole emphasizes optimization by natural selection in large populations with additive effects of genes, the other interaction of genes and interaction of natural selection and genetic drift. ►

I was surprised, too, to find how much of Wright's time during the 1930s and 1940s was taken up analysing Theodosius Dobzhansky's data for his famous series of papers on the "Genetics of Natural Populations", and how dependent Dobzhansky was on Wright's suggestions and analyses. Provine's treatment is extensive and illuminating, and indicates the underlying objectives of many of Dobzhansky's papers. Indeed, Provine's book will be an excellent starting point for a study of Dobzhansky's papers; through it the new student in evolutionary biology may be able to appreciate that genetic variation was once much more difficult to find and measure than it is today.

As the title indicates, *Sewall Wright and Evolutionary Biology* concentrates upon Wright's views and influence on evolutionary biology, and is not a general account of Wright's life or science. Although Wright's work on mathematical genetics is mentioned, Provine devotes little space to it. A symptom of this is the omission of the five essential "systems of mating" papers of 1921 from the collection of papers. On the mathematical theory Provine makes a few false steps. After pointing out that there are three logically distinct versions of Wright's concept of a fitness surface, he argues forcefully that these are logically unconnected, that Wright has spread confusion by using two of the notions interchangeably, and that population geneticists mistakenly assume that the two notions are related. Provine is mistaken in this: you really *can* derive the gene frequency adaptive surface from the genotype adaptive surface by taking a weighted average.

Wright's personal life is treated, but only rather briefly and circumspectly. Many population geneticists have Sewall Wright anecdotes, most of which do not appear in Provine's book. I did not find my favourites: "Sewall Wright And The Coffee Cup", "Sewall Wright Penniless In Berkeley" and "Sewall Wright Lectured On Genetic Drift". To his credit Provine does mention the most famous of all, noting both Wright's denial and the numerous people who claim to have seen him absent-mindedly erase a blackboard with one of his guinea pigs.

Twenty years ago I happened to be the person who first suggested to Provine that he should interview Sewall Wright (Provine was surprised to hear that Wright was still alive). Now I'm glad I did, because the result is a remarkable book. As we face the interesting possibility that Wright will be the keynote speaker at his own centennial celebration, his influence on evolutionary biology is greater than it has ever been. We will be listening to that speech with more than just polite attention. □

Joe Felsenstein is a Professor in the Department of Genetics, University of Washington, Seattle, Washington 98195, USA.

Redouté requiem

Sandra Raphael

Wonders of Creation: Natural History Drawings in the British Library. By Ray Desmond. *The British Library*: 1986. Pp.240. £25.

A Redouté Treasury: 468 Watercolours from Les Liliacées of Pierre-Joseph Redouté. Text by Peter and Frances Mallary. *Dent, London/Vendome Press, New York*: 1986. Pp.228. £40, \$50.

Wonders of Creation contains a selection of illustrations of plants and animals from manuscripts and drawings in various departments of the British Library, with a few pictures from printed books added to demonstrate the changes in style and taste described in the text. The choice seems to have been made almost at random in an attempt to include something for everyone, from flowers and trees, to birds, fishes, snakes, terrapins, rhinoceroses, koala bears, beetles, spiders and Stamford Raffles's "hairy-nosed otter". The sources range from mediaeval manuscripts to nineteenth-century drawings. The coloured plates (nearly 50 of them) are at the back of the book, each with its own short description, while the black-and-white illustrations are used to break up a short text that gives a rather pedestrian survey of the development of illustration as an aid to the descriptive work of naturalists.

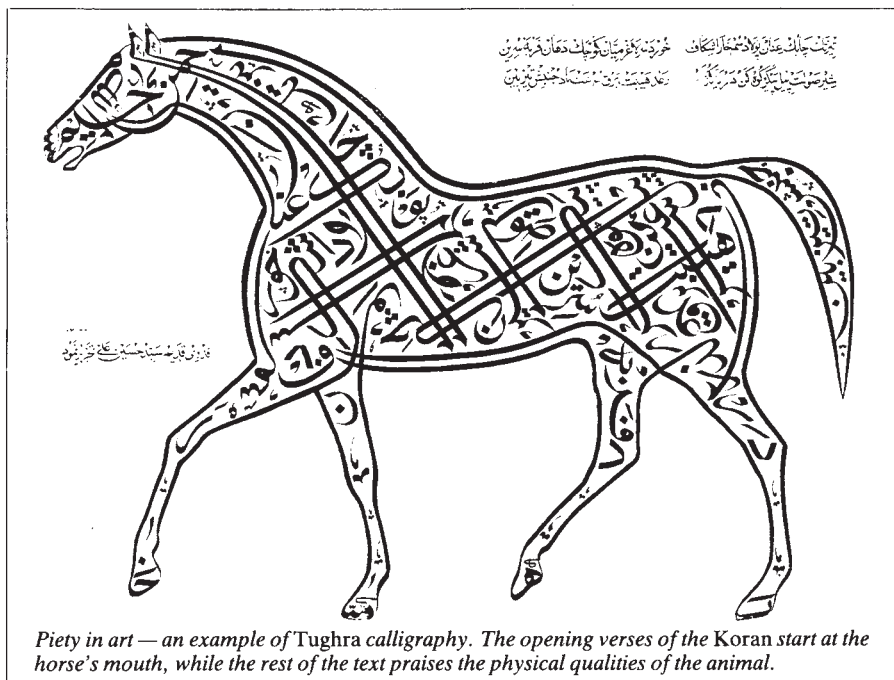
The British Library collections in this field are quite rich, but it seems perverse to isolate them from the complementary and even richer hoards in the Natural History Museum in Kensington and the Print Room of the British Museum, still under the same roof as the British Library. In

spite of the book's restrictions, the beauty of many of the pictures, oriental or European, will surely please browsers.

On 20 November 1985 Sotheby's in New York sold for \$5.5 million the Empress Josephine's copy of Pierre-Joseph Redouté's *Les Liliacées* (1802–1816), 16 volumes containing 468 original watercolours on vellum as well as the printed text. *A Redouté Treasury* is a revised version of the sale catalogue, with colour photographs of the drawings of plants of the lily family and related groups rearranged and supplemented by a longer introductory essay and botanical notes.

Although *Les Liliacées* was sold as a single lot, it has since been dismantled by its buyer, a "breaker" who specializes in taking apart colour-plate books and selling their illustrations piecemeal. This process is sad enough when its subject is a relatively ordinary book, but when it is applied to one that is unique it becomes a bibliographical scandal. Redouté's drawings for *Les Liliacées*, possibly his masterpiece, were bought by his patron Josephine, whose garden at Malmaison provided many models for the artist. The drawings were inherited by Josephine's son, and all but 19 of them were bought by an American dealer, Edward Weyhe, in Zurich in 1935. Now they have been sold again. The drawings have been taken out of their fine bindings, so that this example of a great artist's greatest work has now been scattered and will never again be seen whole. *A Redouté Treasury* may be seen as a requiem for a French national treasure, now ruined. □

Sandra Raphael, 18 Argyle Street, Oxford OX4 1SS, UK, was the natural history editor of the new Supplement to the Oxford English Dictionary, the last volume of which was published earlier this year.



Piety in art — an example of Tughra calligraphy. The opening verses of the Koran start at the horse's mouth, while the rest of the text praises the physical qualities of the animal.

The scorbutic story

John Rivers

The History of Scurvy and Vitamin C. By Kenneth J. Carpenter. Cambridge University Press: 1986. Pp.288. £27.50, \$39.50.

JOHN Le Carré might have told this story better, but I doubt that he could have produced a better book. Like Kenneth Carpenter, Le Carré could have told the tortuous tale of 400 years of medical misunderstanding of scurvy, leading along false trails, down blind alleys and round and round in circles. And Le Carré would if anything have made the reader more aware that, because scurvy was a disease that particularly afflicted armies and especially navies, we are dealing not only with medicine but with grand strategy. The real difference in treatment would come in the last chapter where, the history of scurvy having become the story of the discovery of vitamin C, the author presents a scientific retrospect in which centuries of confusion are put beneath the spotlight of modern knowledge. If Le Carré had unleashed Smiley on these problems, the dénouement would have been far more slick.

Professor Carpenter, however, has been too honest and thorough in his analysis to pretend that all the questions are now resolved. Some, of course, are: the mystery of why the Esquimaux didn't get scurvy, despite eating no vegetables and fruit, is explained by the ascorbic acid content of fresh raw meat. But other problems remain perplexing even today. Why, for example, were so many cases of shipboard scurvy apparently cured by the once-favoured mixture of vinegar and potassium nitrate? The mixture contains no vitamin C; so was it lies, self deception, a placebo effect on an unprecedented scale or something we don't yet understand. The problem is left hanging.

Some certainties fall apart beneath Carpenter's critical analysis. James Lind, for example, doesn't seem anywhere near so epoch-making when Carpenter has finished with him. On the bicentennial of the publication of *A Treatise on Scurvy*, Lind was beatified by scientists as the inventor of the clinical trial and the first truly modern nutritionist. But the Lind to whom Professor Carpenter draws our attention ends up more human, someone who despite his studies on lemon juice as an antiscorbutic believed that cold and damp air actually caused scurvy. By the end of his life, it seems, Lind was a deeply pessimistic man, doubting his own observations which he was unable to replicate in studies at Haslar hospital, where even the controls mysteriously recovered, and concluding in the third edition of his *Treatise*

that "enlarged experience must ever evince the fallacy of all positive assertions in the healing art".

Captain James Cook, too, is seen in a more perceptive light. Cook's circumnavigation of the world was free from scurvy not simply because he used Lind's prescribed lemon juice. He took almost every antiscorbutic that anyone recommended, with the result that the trip provided no test of Lind's hypothesis. Indeed, says Carpenter, Cook may have delayed the general introduction of lemon juice as a prophylactic because every theorist drew sustenance for their pet idea from his provisioning strategy; the award of the Royal Society's Copley medal to Cook, for his success in avoiding scurvy, seems to have been supported by Sir John Pringle, the President of the Royal Society, because he regarded Cook's triumph as evidence in favour of his favourite hypothesis that malted grain was the preventative of choice.

Such re-evaluations are the result of careful research on Professor Carpenter's part, not idle iconoclasm. Indeed, in springing to the defence of the Admiralty, Carpenter attacks the iconoclasm of others. It is often held that the continuance of scurvy in the British navies during the nineteenth century was the result of the Admiralty's penny-pinching decision to abandon lemon juice in favour of lime juice from the West Indian colonies which, though cheaper, had no vitamin C activity. The truth is not so simple. Not only are limes a good source of vitamin C, but the words lemon and lime were used interchangeably and it is not always possible to be sure what was being done. Carpenter suggests that the ineffectiveness of lime juice may have been due to catalysis of the oxidation of the ascorbic acid by copper in the stills and vessels used to concentrate the juice, compounded by overlong storage. The Admiralty was doing its best; it was the scientists who were wrong.

This is an extremely detailed book, and at times I wondered whether Carpenter should not have begun with the modern knowledge of vitamin C to aid the reader in interpreting the sheer mass of erroneous evidence and theory that runs down the centuries: a sort of answers first, questions afterwards strategy. But the lack of certainty in the final clarification shows that we don't yet have all the answers, and some questions must just remain. That alone makes the book worthwhile, as well as delightful to read, and a volume that will inform and entertain both historians and scientists — and, indeed, anyone who likes a good mystery story, even if it isn't written by Le Carré. □

John Rivers is a Senior Lecturer in the Department of Human Nutrition, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK.



INSTITUTE OF TERRESTRIAL ECOLOGY

Landscape changes in Britain

C.J. Barr, C.B. Benefield, R.G.H. Bunce, H.A. Ridsdale & M. Whittaker

A largely visual presentation of results from 2 landscape surveys of Great Britain.

28pp, 0904282953 £3.00

Pollution in Cumbria

P. Ineson

Proceedings of a symposium concerned with acid rain, radioactivity, industrial and agricultural pollution.

92pp, 0904282961 £5.75

Coast dune management guide

D.S. Ranwell & R. Boar

How to use vegetation for coast protection activities, with particular reference to sand dunes.

105pp, 0904282937 £6.00

Distribution and status of bats in Europe

R.E. Stebbings & F. Griffith

31 species are described under the following headings: distribution, habitat, population, threats, conservation measures, bibliography.

142pp, 0904282945 £5.00

Full list from:
Publications Section
Merlewood Research Station
Grange-over-Sands
Cumbria LA11 6JU, UK

Many earlier titles now reduced in price.



Inside development

John C. Gerhart

Molecular Biology of Development. Volume L, Cold Spring Harbor Symposia on Quantitative Biology. Cold Spring Harbor Laboratory: 1986. Pp. 920. Hbk \$140; pbk \$70.

AT THE very least, this is yet another outstanding volume to arise from the annual Cold Spring Harbor Symposium, now in its fiftieth year. The papers — over 100 of them — come from leading molecular biological and genetic laboratories; they cover the subjects of nucleo-cytoplasmic interactions, cell lineage, segmentation, homoeotic mutants, homoeoboxes, tissue-specific gene expression, sex determination, cellular differentiation and developmental neurobiology. There is probably no book on the molecular biology of development to match this one, not only in the comprehensiveness of the coverage, but also in the care of preparation of individual papers, an indication of the authors' high regard for the occasion of a Cold Spring Harbor meeting. Most of the contributions are no-nonsense research accounts which assume a technical background and discuss only the immediate implications of results. Broader orienta-

tion is available by way of a general introduction by John Gurdon, and a summary of symposium highlights by Gerald Rubin. In this rapidly changing field, the papers cannot be expected to remain current for very long; the volume should be studied forthwith.

What is the molecular biology of development, as revealed in this volume? Joe Sambrook writes in his foreword:

Developmental biologists are an audacious breed. Their ambition is to provide an intellectually satisfying account of the forces that guide development of multicellular organisms around the circle from fertilized egg to adult to gamete. These processes, however, lie close to the border beyond which genetics may not be feasible nor necessarily useful; furthermore, meaningful biochemical studies pose extraordinary challenges because of the small amounts of material that are available and the general rapidity of events in the early embryo. In the face of these difficulties it is not surprising that for many years developmental biology was essentially a branch of anatomy, in that it was exclusively descriptive in nature. Accurate and elegant though these descriptions were, they led to an understanding of developmental processes less frequently than to an amazement at their beauty.

In recent years, however, developmental biology has undergone a dramatic change and has matured from a descriptive to an analytical science. There is no doubt in my mind that this change stems almost entirely from two technical advances — the ability to use molecular cloning to isolate and characterize mutant and wild-type versions of genes that control or are expressed at specific developmental stages and second, the ability to generate transgenic organisms in which the expression of the introduced gene(s) is correct both spatially and temporally. In consequence, the developmental biologist has not only the capacity to describe but also now to analyze and influence the events that guide a fertilized egg to its destiny.

In the book, it seemed to me that the excellent description by I. Herskowitz of the mating-type locus in yeast approaches most closely the kind of "intellectually satisfying account" now aspired to by molecular biologists studying development (for example those working on sex determination, segmentation and homoeosis in *Drosophila*). Herskowitz indicates (1) how certain master regulatory genes encode DNA-binding proteins which directly control the expression of batteries of cell-type-specific genes, and (2) how these master regulatory genes are themselves regulated by products of the *HO* gene, which is activated within mother cells by the products of yet other genes of the *SWI* family. Finally he suggests (p. 572),

returning to the egg, we can imagine that regulatory proteins such as the *SWI* products become asymmetrically distributed to sister blastomeres. They activate expression of genes analogous to *HO* that govern activity of a regulatory locus (analogous to *MAT*) that ultimately triggers cell specialization.

Several contributors trace similar

thoughts to those of T. H. Morgan (*Embryology and Genetics*, 1934, p. 10):

The initial differences in the protoplasmic regions [of the egg] may be supposed to affect the activity of genes. The genes will then in turn affect the protoplasm, which will start a new series of reciprocal interactions. In this way we can picture to ourselves the gradual elaboration and differentiation of the various regions of the embryo.

Thus, in scanning this volume, I sensed fair agreement among participants that new information on genetic regulatory hierarchies is compatible with the long-appreciated phenomenon of cytoplasmic localization (an *intracellular*, lineage-specific phenomenon) and can be grafted on to it to give an intellectually satisfying account now identifiable as the molecular biology of development.

Is there more? As John Gurdon points out, vertebrate development is thought to depend primarily on arborizing cascades of ever more local inductive cell-cell interactions. The highly regulative nature of vertebrate embryos indicates the ongoing ability of individual cells to adapt their paths of development to *intercellular* conditions. Imaginal disc regeneration and transdetermination also reveal the importance of continuous cell-cell interactions in invertebrate (*Drosophila*) development. These complex, multiple interactions, while mentioned in the volume, have an uncertain relationship to the genetic control of cell differentiation and to lineage-restricted cytoplasmic factors. If embryonic cells produce and respond to a great variety of *intercellular* contexts, and if a cell's particular path of development (including its changing profile of gene expression) is highly adaptable to that context, then regulatory genes would themselves be subject to extensive regulation, and would seem to be rather less masterful in their control of development, even though they would function to control other genes and play an important inertial role in connecting the cell's past to its present. In organisms making use of cell-cell interactions to select developmental paths, gene-regulating genes might be elements of much larger intercellular regulatory circuits having many components, many inputs, multiple states and adaptability. (It's one thing to say that a gene controls other genes; it's another to say that a gene controls development.)

If this is the prospect, then with the help and stimulus of the exciting new analytical techniques, and with genetic dissection, we can look forward to a long period of amazement at the beauty of our enlarging description of development, as the molecular biology of development turns outwards from the nucleus. □

John C. Gerhart is a Professor in the Department of Molecular Biology, University of California, Berkeley, California 94720, USA.

NUCLEAR POWER AND ITS ENVIRONMENTAL EFFECTS

For the Layman, Nuclear Technician, Manager

Places the effects of possible nuclear radiation in perspective. Meets the need for a better understanding of radiation and its consequences especially after the TMI accident.

All vital aspects of nuclear power generation are considered, from fundamental principles to plant licensing, from uranium mining to disposal of waste. The safe operation of nuclear power plants, to prevent the escape of dangerous amounts of materials in case of an accident, is fully covered.

By Samuel Glasstone
Walter H. Jordan
American Nuclear Society
555 No. Kensington Ave.
La Grange Park, IL 60525
USA

408 pages
ORDER TODAY
\$18.95 Softbound
\$26.95 Hardbound

The source for the latest knowledge in Nuclear

Feeling for nature

Robert M. May

Gilbert White: A Biography of the Author of *The Natural History of Selborne*. By Richard Mabey. *Century Hutchinson, London:1986. Pp.239. £14.95.*

ONLY three books in the English language have gone through more editions than Gilbert White's *The Natural History of Selborne*. This claim leaps out from the blurb on the back of the 1977 Penguin edition, and has left me wondering what ranks third behind the Bible and Shakespeare; probably *Pilgrim's Progress*. The separate editions of *Selborne*, over 200 of them, have appeared at a steady rate since the book was first published in 1789 (note the impending bicentenary), forming an interesting record of changing tastes in book design and changing visions of the countryside as revealed in the illustrations.

On first acquaintance, *Selborne* is a surprising book to have evoked such unwavering affection. As Mabey observes, it is

the kind of book presented on prize-giving days, ... a work, in all senses, of the old school. Even when I eventually came to read it, I cannot say my opinion changed dramatically. I could not cope at first with its rambling disorder, its sudden plunges into thickets of taxonomic Latin, and, for a while, I failed to notice the feeling behind the often dispassionate prose.

Selborne is organized as a collection of letters, similar to many other works about local natural history by enthusiastic parsons. It endures while these other books ornament the libraries in stately homes because beneath its unassuming surface is a fully contemporary appreciation — quite new in its own time — of plants and animals as more than curiosities, but as parts of a community, interacting among themselves, with their environment and with humans. White raises general questions (such as how populations are regulated), but always against a background of keenly observed detail (the breeding success of swifts, the singing of crickets, or the way fly-catchers keep their fledglings cool in summer by fanning their wings above the nest).

Richard Mabey has produced the first biography of Gilbert White for over 80 years. The difficulty is that not a lot is known about White, except that he lived an uneventful bachelor life in the parish and rarely left it (as Richard Boston observed, reviewing the book in *The Guardian*, how else could White have described the South Downs as "that chain of majestic mountains"). Mabey's book is more in the nature of an introduction to, and appreciation of, *Selborne*, than a biography in the strict sense. At such, it

does a fine job of relating Gilbert White and his book to the village of Selborne as it was then and is now. Despite his observations, White never fully gave up the contemporary notion that swifts and swallows disappeared in winter because they hibernated; Mabey empathizes, suggesting that White simply could not imagine why anyone could want to leave Selborne.

Mabey explains that "it was seeing Selborne itself, the source of White's inspiration, that changed my view [quoted above] decisively, and helped me to understand what the book was truly

about". The patterns of field and woodland around the village have changed little since White's time, and the village itself conducts its Gilbert White trade in a low-key way (the stained glass memorial window in the church, with White as Saint Francis among his birds, is alone worth a visit: see the cover illustration of *Nature*, 22 December 1983). Visit Selborne and read Mabey. Then read *The Natural History of Selborne*. □

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, Princeton, New Jersey 08544, USA.

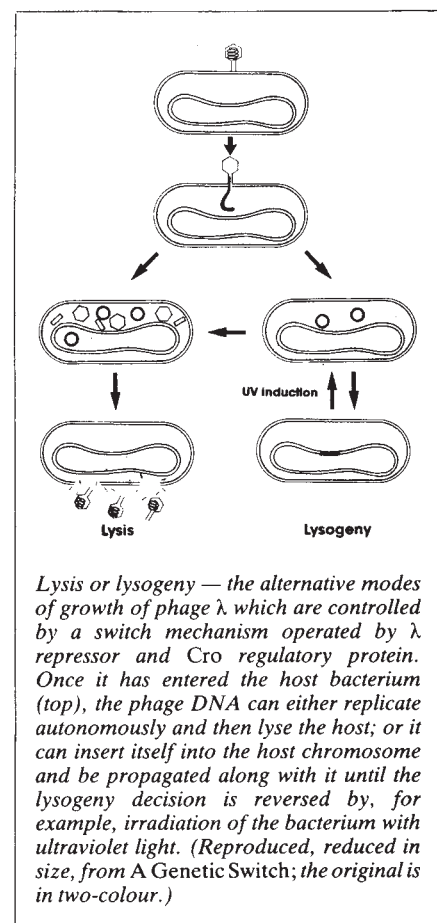
In the groove

Sydney Brenner

A Genetic Switch: Gene Control and Phage λ . By Mark Ptashne. *Cell Press, Cambridge, Massachusetts/Blackwell Scientific: 1986. Pp.128. Pbk \$16.95, £14.50.*

MAX Delbrück did not believe in lysogeny. He thought that the persistence of bacteriophages in some bacterial cultures was due to cryptic infection and he dismissed it as a fringe phenomenon. Everybody he influenced therefore worked on virulent bacteriophages, especially T4, and their work led to one branch of molecular genetics and to the discovery of the fine structure of the gene, the general nature of the genetic code and messenger RNA. Fortunately, Delbrück's influence was not totally pervasive and in Paris, where there was a long tradition of research on lysogenic bacteria, the work of Lwoff and of Jacob and Wollman on bacteriophage λ revealed that lysogeny is in fact due to a gene regulatory mechanism that allows the phage DNA to remain silent in the bacterial chromosome until activated by, for example, ultraviolet light. Their major step was the realization that the properties of the phage mutants on which they worked were isomorphous with those affecting the expression of β -galactosidase, and that in both cases gene regulation was achieved by the negative effect of repression.

Mark Ptashne isolated the λ repressor in 1967 and has spent nearly 20 years in pursuing its molecular structure and mode of action. This little book is an account of his research, focused especially on how the λ repressor and *Cro*, another regulatory protein, form a switch by their interaction with an operator, O_R . In one state, corresponding to stable prophage, the binding of λ repressor inhibits the rightward transcription from the promoter site P_R and stimulates its own production by transcription in the direction from another promoter, P_{RM} . When the repressor is removed by proteolysis induced by ultraviolet light, rightward transcription is



turned on and *Cro* is produced. The binding of *Cro* to O_R prevents any further synthesis at P_{RM} , thus turning off repressor, irreversibly throwing the switch and sending the phage down the lytic pathway. Ptashne shows how these events are precisely conditioned by the mode of binding of repressor and *Cro* to the three-component sequence of the operator. The structure of these proteins is known and both recognize the major groove of DNA through a characteristic α -helix-bend- α -helix, a motif now known to be present in a large number of DNA-binding proteins.

This work is a paradigm for the mechanisms of regulation but Ptashne would like to go further and sees the λ switch as the model for switching between develop-

mental pathways in higher organisms. In a general sense this is true because, logically, the λ system appropriately combines positive and negative feedback in a simple way to implement a switch; but this is not necessarily the only way it can be done. Years ago Novick performed a most illuminating experiment. Normally the galactosidase of *Escherichia coli* is repressed, but it can be released from repression (induced) in the presence of lactose or an analogue thereof. Novick briefly exposed *E. coli* cultures to a high concentration of an inducer which was then diluted a thousandfold to provide a maintenance level. In a fraction of the cells, enough of the inducer got in to derepress the genes; the rest remained uninduced. The resulting mixed population maintained itself indefinitely as long as the low level of inducer was present, because induction entails the production of a permease that increases the uptake of the inducer so that even at low levels enough can enter the cells to maintain the induced state. Thus it could be said that *E. coli* had "differentiated" into two cell types with heritable states. This result depends on the properties of the permease: only cells with permease can take up inducer at the low maintenance level and only such cells can have permease induced; transient induction at high levels simply initiates the process in part of the population. Here the logical elements of positive and negative feedback are implemented in another way and it is arguably a better paradigm for some developmental processes than the λ switch.

The analogy between the λ switch and developmental processes should not be carried too far. In discussing the difference between "commitment" and "differentiation" of cells, Ptashne writes

By analogy, a λ -lysogen is indistinguishable from an uninfected bacterium by casual inspection, but the former and not the latter is 'committed' to lyse and release phage when it encounters an inducing signal in the environment.

However, he quickly realizes that this is not a good analogy because the difference between the bacteria is a genetic one and he continues

The lysogen, of course, has genes that are lacking in the non-lysogen, but it is not difficult to imagine analogous forms of commitment involving only differential gene expression. To take a simple example, the regulatory proteins of a cell might turn on a gene encoding a hormone receptor.

Do we hear a faint echo of the permease paradigm?

But this is quite a minor criticism of a clear exposition of a series of triumphant experiments and insights that have brought us close to the molecular heart of the matter of genetic regulation. $\square$

Sydney Brenner is Head of the Medical Research Council Molecular Genetics Unit, Hills Road, Cambridge CB2 2QH, UK.

Eastern originality

Christopher Cullen

China: Land of Discovery and Invention. By Robert K. G. Temple. Patrick Stephens, Wellingborough, UK: 1986. Pp.248. £12.95.*

READERS of this journal will need no more than a brief reminder of the importance of the work of Joseph Needham. With the help of a large group of collaborators, over the past 40 years he has been engaged in a great project of research and publication on the history of science, technology and medicine in pre-modern China. The results are appearing in the pages of *Science and Civilisation in China*, the seven volumes of which are being published by Cambridge University Press as some 30 separate tomes (about half of them have appeared to date). This work is widely regarded as the most influential British contribution to historical scholarship in the present century, and it has certainly completely changed the perspective of many well-informed Westerners on the origin of important aspects of modern science.

As the weight of his research has accumulated, Needham has had to move far from his original aim of writing a short book for the general reader. Robert Temple's book, which is compiled almost entirely on the basis of material already published or awaiting publication in Needham's volumes, is intended to meet this need. Temple has divided his account into 11 broad subject areas, such as "Agriculture, Medicine and Health", "Magnetism" and "Warfare". Under these he has

grouped a hundred short, illustrated essays on particular topics. Temple is clearly no mere hack compiler of coffee-table books: his writing is clear and vigorous, and he succeeds in communicating his obvious enthusiasm for the subject. It is hard to imagine anyone with a general interest in the history of science failing to find a great deal in this book to intrigue and delight.

There are some flaws. Temple makes no systematic attempt at writing history; the book is a series of vignettes on "the wonders of Chinese science", and as such does not convey any sense of the social and philosophical background from which Chinese science, technology and medicine arose and against which they functioned. The choice of topics and the balance between them sometimes seems odd: the section on engineering devotes more space to the novelties known as "spouting bowls" than to cast iron. Some of the statements are out of date or just plain wrong, and the one-page list of "further reading" is so incomplete that it does not even mention Colin Ronan's readable and reasonably priced abridgement of Needham's work, also published by Cambridge University Press. On the whole, however, Temple deserves the gratitude of his readers for having provided a ready means of access to a fascinating subject. $\square$

Christopher Cullen, 86 Mount View Road, London N4 4JX, UK, is a former Research Fellow of Clare Hall, Cambridge, and has published mainly in the field of ancient Chinese astronomy and mathematics.

*To be published in the United States in January by Simon & Schuster, under the title *The Genius of China: 3000 Years of Science, Discovery, and Invention*. Price will be \$19.95.



Sowing sophistication — spring wheat being sown by seed drill in northern China. Seed drills, which are much less wasteful than broadcasting seed by hand, were available in China from the second century BC. It was almost 2,000 years later that they came into use in Europe.

Deception beyond measure

C.R.B. Joyce

False Prophets: Fraud and Error in Science and Medicine. By Alexander Kohn. Basil Blackwell: 1986. Pp.221. £14.95. To be published in the United States in January 1987, \$24.95.

AS THIS book and its chief rival, William Broad and Nicholas Wade's *Betrayers of the Truth** both remind us, Charles Babbage distinguished three unorthodox ways of dealing with data. These he called, in increasing order of impropriety, trimming, cooking and forging. Are such behaviours more common today than they were in Babbage's time? In his very first sentence Alexander Kohn states that the exponential growth of science has been accompanied by the "appearance" of deviant scientists, as if a rectilinear growth would not have given rise to the problem. It is not clear whether he thinks that there has been an increase in the number of such individuals, as Broad and Wade imply; and, if so, whether this is absolute or proportional. In fact, there have been more attempts than Kohn realizes to answer the question, but it is not an easy one to tackle, requiring as it does methods learnt from the confessional, the KGB and Social Audit. But whether or not the wickedness has increased or is increasing, we are sure that it should be diminished.

When a question arises in the science of science (a respectable discipline, the methods of which have been successfully applied to the growth of science as a whole), one would expect scientists to try to answer it scientifically. Too often, however, they fall back on dogmatic or wistful assertions of belief. If public attitudes to science were still as starchy-eyed as they were even 25 years ago, such piety might suffice. But reassurances that there is nothing to worry about — because the scientific process is self-corrective, and that mistakes, deliberate or accidental, will be found out sooner rather than later — are no longer enough. Such reasoning belongs to the "tap-dancing and dinosaurs" class ("I'm tap-dancing to keep the dinosaurs away!"). "But there haven't been any dinosaurs for years!". "There! You see?"). That the only examples of fraud that have been discovered are the ones that have been discovered is not like-

*Published in 1983 by Simon & Schuster, and reviewed in *Nature* 302, 774 (1983).

ly to convince the critics of science, and should not satisfy scientists either. Even Popper considers only that errors are "fairly often, in time, corrected".

How much time is needed for the purpose? Seventeen years or so between deed and discovery in the case of Cyril Burt, or 40 for the Piltdown hoax or Gregor Mendel? There are even more important questions to ask about the ethics of scientific discovery than this, but they are seldom examined systematically. One obvious framework is: What is scientific fraud (definitions)? How much of it is there (estimates, including trends)? Why does it occur (reasons — social as well as individual)? Does it matter anyway (costs — and possible benefits: see below)? And what is being and should (or can) be done about it?

TRICKS OF THE TESTING TRADE

Pivate testing laboratories live and die by the positive results generated on products they test for safety. Your future as a laboratory owner or administrator is in serious jeopardy if you find too many products unsafe. But fortunately, manufacturers rarely run their own audits on tests contracted out to private labs. That way, they can claim ignorance of wrongdoing if protocol is not followed. So here are some tricks you can use to help that drug, chemical or product meet approval standards.

THE HAPPY ENDING. Bureaucrats rarely seem to have time to read more than summaries and conclusions. So you can do your manufacturing client a great service by having the conclusion bear minimal relationship to your findings, particularly if the findings are negative. Remember that statistics are in the service of those that use them.

THE SHRED. Keep a paper shredder near your records at all times. If you hear that a government inspector or auditor is about to visit, use the shredder quickly to rid your files of embarrassing or incriminating data. For fast shredding techniques, call Industrial Bio-Test Laboratories.

THE BENCHMARK BOOGIE. If you have to create new benchmarks to subvert

hiring untrained technicians, you can always use them as an excuse when the scientific method is inadvertently abused or abandoned. You can contact almost any testing lab in the country in your search for undereducated, unqualified and incompetent staff. Move your staff around the lab constantly throughout experiments. If someone is leaving your firm, have that person sign all the benchmarks. That way the "responsible party" will be absent should an inspection ensue.

Kkeep a scalpel in your lab, a shredder near your files, and don't tell the Feds your data caught fire.

THE SPEEDUP (OR THE TIME-LINE FUDGE). Many safety tests require long exposure of the product to hostile environments like extreme heat, dryness, cold or wetness. Take semiconductors, for

on the creative use of the scalpel contact the G. D. Searle Company and ask about the tests it performed on Aldactone, a diuretic prescribed for high blood pressure and fluid retention. Or write to Syntex Corporation and ask about the research it contracted for on the antihistatic drug Naproxen.

THE POSTHUMOUS APPEARANCE. Should laboratory animals die, replace them with live, healthy relatives. Again, contact G. D. Searle and Industrial Bio-Test Laboratories in Chicago to learn how some of their rats, monkeys, rabbits and dogs, which died during tests, later appeared alive in test data. Animals can also be recorded as gaining weight and recovering from tumors after their death. Use your benchmark creatively.

THE DATA MESSAGE. Also known as "Graphiting," data massaging is probably the most common fudging technique used in labs today. Motivated by what scientists euphemistically call "intentional bias," it is really the easiest way to get the results you want once the experiment is completed. There is no one way to massage data. Just take those long computer reports home with you, spread them out on the kitchen table and gently apply rubber and graphite.

SOME EXTRA ADVICE: If the validity of your study is questioned, do not under any

on scientific fraud would begin, instead of ending, at that point. Given one society (Britain) in which the world's largest advertising agency can switch from helping a national health campaign to promoting the very product whose use it had previously sought to diminish; another (the United States) in which 25,000 unnecessary cardiac pacemakers are inserted per year; a third (the Soviet Union) in which . . . The list can be continued to infinity.

One should not, of course, criticize an author for writing the book he wanted to write; but it is fair to suggest that Kohn has not carried out his own expressed intention. His sub-title is *Fraud and Error in Science and Medicine*, less alarmist than Broad and Wade's *Fraud and Deceit in the Halls of Science* more open-minded

too. It also allows Kohn to mention Mendel and Newton more properly in his book than Broad and Wade in theirs. But though with these references, as well as elsewhere (for example, in the attractive dust-jacket of Wyck's "Alchemist") Kohn implies that error may sometimes be only an idea whose time has not yet come, the theme nowhere receives the discussion its importance requires and the title of the book promises: When is an error not an error but a flash of inspiration? Not only do our hypotheses depend upon our perceptions, but perception is itself a prisoner of hypothesis. Even good scientists have sometimes been so sure that they were right that, in Alan Franklin's words, they "manufactured data . . . until real data became available" — even one of Dar-

see's retracted papers has already been de-retracted!

Perhaps Kohn fails to deliver because, understandably, he is as uncertain about the matter as the rest of us. In fact, he is at times so uncertain in his handling of certainty in general — for example, the meaning of correlation, and the relationship of truth to the laws of chance; not to speak of his confusion between experimenter (or cognitive) error and experimental (or statistical) error — that the reader is sometimes driven to doubt his grasp of the matters of which he writes.

According to Stephen Toulmin, "medicine has saved the life of ethics" (*Perspectives in Biology and Medicine* 25, 736–750; 1982), producing "spectacular and irreversible effects on the methods and content of philosophical ethics". It is high time for at least a first instalment towards repayment of the debt. □

C.R.B. Joyce is Head of Project Innovation in the Medical Department of Ciba-Geigy Ltd, CH-4002 Basel, Switzerland.

Of course, Kohn is aware of these problems but, as is common among popularizers of science, he seems to believe that readers prefer a stream of hair-raising stories to a serious discussion of ways to identify and prevent fraud in the future. Thus it would serve little purpose to quote the table of contents; the divisions between subject areas often seem to be arbitrary and the chapter titles usually promise more than the text delivers. Even though his final chapter is entitled "What Can Be Done about Fraud in Science?", and might have been expected to bring together and add to suggestions scattered about the book, it is only three pages long and one of them is taken up by introducing two further anecdotes warning would-be whistle-blowers that theirs is a very chancy and perhaps even dangerous activity.

The most relevant point is made in the very last paragraph of all: "To understand deceit in science, one has to study first the ethics of the society within which the scientists work and live . . .". A useful book

The chips are now down?

Steve Woolgar

Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer. By Hubert L. Dreyfus and Stuart E. Dreyfus. Basil Blackwell/Free Press: 1986. Pp.231. £15, \$16.85.

THE most important and exciting thing about artificial intelligence (AI) is that it provides the possibility of a critical test of fundamental assumptions about the nature of knowledge and reason. These assumptions have gripped Western thought for 2,000 years and they underpin ingrained ideas about the character of man, his abilities and potential. If the project to devise artificial reason is adjudged a success, it will vindicate the objectivist and rationalist philosophies which are thought to have guided the rise of science and technology; its failure will buttress those phenomenological and interpretivist philosophies which argue that calculation and measurement are largely irrelevant to man's capacity for intelligent behaviour.

Hubert Dreyfus recognizes what is at stake. His *What Computers Can't Do*, first published in 1972 and revised in 1979, was a compellingly sophisticated analysis of the philosophical assumptions which underlie AI. That book provided an overview of developments in AI from work on cognitive simulation in the 1950s to early attempts at knowledge engineering in the 1970s. *Mind Over Machine*, written with his brother (a computer specialist), extends the analysis to recent work, particularly on expert systems. The book's appearance is timely, in view of the sudden and massive increase in financial support for AI, especially from military sources.

Will these people collect on their expensive bet? The Dreyfuses say no. Their central thesis is that intelligent human behaviour (interpretation, understanding, recognition, puzzle solving, rule application and so on) is simply not amenable to formalization in the way required by digital computers.

The authors propose a model comprising five stages of skill acquisition, ranging from novice to expert. Progression beyond the third stage ("competence") requires a qualitatively different approach to reasoning. Expertise, in their view, involves the accumulation of experience, the unconscious holistic recognition of the familiarity of new situations and the critical use of intuition and "know-how". This is beyond the ability of systems designed on principles of calculation from discrete facts and codified rules. The apparent

success enjoyed by some artificial devices occurs in restricted (artificial) areas which require competence rather than intelligence. This is not to deny the usefulness of some AI systems as aids, but such systems are properly termed "competent" rather than "expert". The five-stage skill hierarchy is used to discern the limits of computer use both in education and in managerial decision-making. Only in the rare instances, where researchers abandon the mechanistic models of human skill, do we find hope for a better approximation to human expertise.

The anti-mechanistic view of human skill derives from the Wittgensteinian argument that the determination of interpretation, understanding, recognition and so on is entirely construed in a social context and can never sensibly occur in isolation.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

tion from it. But the Dreyfuses' hierarchy of skills is strikingly individualistic and mentalistic. For example, they characterize expertise in terms of inner processes: "the expert driver becomes one with his car, and he experiences himself simply as driving, rather than as driving a car . . ." (p.30). Given the philosophical orientation of their critique, there is surprisingly little acknowledgement of the way in which such experiences are meaningful only as socially sanctioned and publicly accountable descriptions. Although the authors insist that their aim is not to mystify, their treatment of skills as inner processes and experiences provides too easy a target for the rationalists. Similarly, it is surely offering too much of a hostage to fortune to continue to speak of the brain and the mind as receptacles of intelligence. The Dreyfuses intend these curious and outdated concepts as a sanctuary for unformalizable human skills. But as others such as Jeff Coulter have shown, notions of "brain" and "mind" are social constructions which tend to support the mechanistic view of human behaviour.

The Dreyfuses are especially impressive in demonstrating that despite changes in emphasis and direction (for example, the shift from attempts to fashion general,

problem-solving devices to those able to make inferences within a highly specialized knowledge domain), workers in AI repeatedly face the same fundamental conceptual problems. For the Dreyfuses, AI's failures are the direct result of such irresolvable problems. For many AI enthusiasts, however, these are merely temporary, technical and surmountable difficulties. For example, whereas the Dreyfuses say that work on "expert" systems shows it is wrong to assume that human experts are using heuristic rules, proponents of AI interpret the research as merely indicating some snags in eliciting these rules. (It is notable that most previous responses to Hubert Dreyfus's ideas have preferred to contest the alleged failure of particular systems, rather than address the conceptual problems he raises.)

Early restricted successes gave rise to unwarranted optimism about the future of AI and, for the Dreyfuses, this optimism partly accounts for the current inability of the AI community to recognize failure. But the Dreyfuses also suggest that the stakes are now so high that failure is actively concealed. They thus allege deliberate cover-up by the AI establishment, describing in the book their own treatment and the manipulation of TV appearances by the "artificial intelligentsia", and criticizing the "unscientific" attitude of AI workers. (The dust-jacket speaks of "AI's record of barefaced public deception".)

This recourse to a kind of conspiracy theory is something of a distraction. For a full appreciation of the dynamics of the AI enterprise, we need to look carefully at the social attribution of "success" and "failure" in the context of a pervasive commitment to rationalistic models of intelligence. Although the Dreyfuses' conceptual argument provides a good, sceptical springboard from which to launch a social analysis of AI research, we should avoid their supposition that conceptual problems will necessarily give rise to "failures".

Who, in any case, are those who cover up failures and spread false optimism about AI? It is important to be clear that the Dreyfuses identify a relatively small body of researchers. True, these are the most influential and respected champions of AI and are thus a legitimate and necessary target for attack; many are entrepreneurs with a substantial interest in furthering public support for the enterprise. But it might be misleading to accept their proclamations as characteristic of AI as a whole. Searle's distinction between "strong" and "weak" AI is useful here. Whereas proponents of strong AI (who claim, for example, that programs can "understand") are appropriate targets for the Dreyfuses, the more cautious advocates of weak AI (who see computers as,

at best, a tool for aiding our investigations into the nature of intelligence) might agree that it is unreasonable to expect the success of artificial systems beyond the level of competence.

Weak and strong AI are not unrelated, of course. Proponents of the former have undoubtedly benefited from the public persuasiveness of the latter. But the distinction points to a further shortcoming in the Dreyfuses' case. By concentrating upon proponents of strong AI, the arguments give us little insight into what practitioners of AI do. In this book, the Dreyfuses continue their impressive assault in the titanic struggle between phenomenological philosophy and the rationalist assumptions of AI's public guardians. But we are not told whether or how the assumptions of strong AI touch upon the day-to-day business of devising systems, making programs work, evaluating output and so on.

It is interesting to consider the likely fate of the Dreyfuses' thesis. Their case against AI is the strongest and most eloquent yet published, a protest against conceptions of the human being which have held sway since the beginnings of Western thought, and it suggests that one outcome of the AI debate might be a definitive refutation of the rationalist metaphysic. However, the lack of a fully social conception of the practice of machine intelligence is perhaps reflected in their inattention to one of the key lessons of the history of science: "critical tests" are rarely decisive. Evaluations of "success" and "failure" need only generate further redefinitions of AI, not its abandonment. □

Steve Woolgar is a Lecturer in the Department of Human Sciences, Brunel, The University of West London, Uxbridge, Middlesex UB8 3PH, UK. With Bruno Latour he is author of Laboratory Life: The Construction of Scientific Facts, a revised edition of which has recently been published by Princeton University Press.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Wondrous gifts

Peter Bryant

Nature's Gambit: Child Prodigies and the Development of Human Potential. By David Henry Feldman with Lynn T. Goldsmith. *Basic Books:1986. Pp.259. \$19.95.*

EVERY now and then a child turns out to be outstandingly talented in a quite specific way. It is not just that he quickly outstrips all his contemporaries in a particular skill: he also soon becomes a great deal better than most grown-ups at it. Children such as this arouse great interest and wonder. Psychologists in particular have to pay attention to them because at first sight they seem to break most of the rules set out by developmental theories. After all, these theories generally hold that different intellectual skills are connected and that they develop together at roughly the same pace through childhood. How can such theories possibly explain a five-year-old child whose mathematical skills already outstrip those of 95 per cent of the adult population, but who otherwise seems pretty ordinary?

I suspect that this question was the starting point for David Feldman's book, which is a prolonged and detailed study of six different, prodigiously talented American boys and their families, but it is not in the end a question that he answers satisfactorily. Instead he spends most time on another question which is unanswerable, or at any rate cannot be answered from a study of individual child prodigies. His main concern is to find what conditions are needed to produce such children. Feldman's own idea, which is about what he calls the "coincidence process" is neither new nor sophisticated: it is that these

children have extraordinary gifts ("proclivities" is Feldman's term) to start with but that their talents only flourish in the right environment. Prodigies need sensitive and devoted parents and an appropriate social environment, according to Feldman, and he is plainly impressed by the lengths to which some of the parents in his little group were prepared to go to support their children. Sacrificing jobs, commuting long distances, confronting awkward teachers, taking the child from one tutor to the next — these seem to be typical demands for a prodigy's parents, and in this group at least they shouldered the burden quite cheerfully. Often the parents' own talents matched their child's: the boy with an extraordinary gift for music had a musician for a father, and the father of the boy who showed early signs of being a talented writer was a writer himself. In these cases the parents spent a great deal of time teaching their own specific skill to their child.

Feldman's argument that all the children in his group were lucky enough to be born into a suitable environment is convincing, but it does not prove his case that such environments are essential for the emergence of a prodigy. The trouble is that his thesis is just as much about children whom he did not study as about the six prodigies with whom he spent so much time. There must, he argues time and again, be children who could have been prodigies, who had the "proclivities", but who found themselves in the wrong environment: the world must be littered with mute, inglorious Miltons or Cromwells guiltless of their country's blood. But Feldman did not study any such prodigy *manqué*. He does not even demonstrate that they exist, and it is very difficult to see how he could have done so, given his methods.

Making music — the prodigiously talented Mozart, accompanied by his father and sister.

Feldman's description of the lives of the six boys is interesting and frequently poignant: but it is difficult to see what is new here. After all, we already know a great deal about the childhood of erstwhile prodigies. Mozart's childhood, to take one example, is documented as well as any other part of his life. Sartre, to take another, wrote about his early years in great detail and with considerable distaste. One of the most striking features of these examples and the children studied by Feldman is the surprise and interest that the prodigies excited in the adults around them. The children soon become performers and they do so at the insistence of adults. On occasion, adult astonishment at the child's unexpected talents has some odd results. In my view the most interesting part of Feldman's book is his description of the way in which parents or teachers sometimes resort to irrational explanations for a child's extraordinary talents, because they cannot account for them in any other way. Reincarnation is one solution. The child is thought to be so talented because he is simply a new version of a dead genius.

To read Feldman's sympathetic and sensitive accounts of these six children is to feel the same kind of awe which the young Mozart must have inspired in those who heard him play and improvise so effortlessly. Readers of this book will wonder at the phenomena which Feldman describes, but will remain puzzled about the reasons for them and about the implications for developmental psychology. □

Peter Bryant is Watts Professor in the Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.

Knockabout stuff

Derek Ager

The Tenth Muse: The Pursuit of Earth Science. By Ronald B. Parker. *Charles Scribner's Sons*: 1986. Pp. 221. \$15.95.

THE nine official muses were responsible for literature, art and a limited amount of science. They had originally been three in number, but six more were conceived and so it is reasonable to add another to represent geology, the most beautiful and inspiring of the sciences. Professor Parker, however, obviously nominates the subject for its entertainment value rather than for its aesthetic qualities.

We are all interested chiefly in our own ideas, but Professor Parker's egocentricity is particularly evident. He assumes that we have read his previous book, and tells us how his great "buffer" idea came to him and how the audience reacted to it at various lectures. Personally I do not understand the "buffer" terminology which is central to most of Dr Parker's thoughts, and would have thought that a better metaphor for what he discusses (the storage and sudden release of energy) would have been the elastic band, which can be stretched so far before suddenly breaking.

Parker's main interest is in our geological today and yesterday — geysers and landslips, glaciers and ice dams. In "The Urge to Surge", for example, he discusses the movement of glaciers and their episodic behaviour, making an assignation with the reader up in Alaska around the year 2000 when things are due to get moving again.

In an excursion into his other main field, crystallography, Parker points out that this science originated in the West, where symmetry is taken for granted, and not in the East where the Zen philosophy prefers asymmetry, be it in gardens or pottery. He goes on to discuss man and other "non-crystallographic" forms in the world (with a delightful independence of thought, he and his wife built themselves a 14-sided house), and Schechtman's discovery of the impossible, that is to say, five-fold symmetry in crystals, as a salutary warning of the need for his own kind of open-mindedness. As a palaeontologist I should point out that echinoderms came up with five-fold symmetry first, albeit using trigonal calcite crystals.

There is a boyhood trip to Boron, site of the largest borate mine in the world, before an essay on "coral clocks" and organic growth generally. Here Parker regrets that corals, with their daily growth-lines, are rare as fossils compared with "clams", thereby displaying a surprising ignorance of the Palaeozoic part of the record. He obviously prefers the "clams" (including their suitability for feeding the poor gra-

duate students studying them), and shocks the palaeontologist by referring to the brachiopods as their "look-alikes". All of this makes me prefer Parker as a geomorphologist and crystallographer rather than as a palaeontologist-stratigrapher.

So we go on to "Aqua Spectaculars and Other Disasters", ranging from Esther Williams and her symmetrical swimmers in those boring films of the 1930s, and via an extremely tenuous link to tsunamis and catastrophically emptying lakes, such as Lake Bonneville and, of course, J. Harlen Bretz's spectacular Lake Missoula. And so back to the "buffer" phenomenon (or better perhaps the knicker elastic of the same vintage as Professor Parker's films) in the building and release of energy in eruptions ranging from Mount St Helen's to Santorini.

There is, too, a general discussion of fossil "scat" (excreta to British readers), including description of a remarkable con-

centration of three billion tons of apatite in the Kola Peninsula which *might* have been the result of a monstrous ball of dung hitting the Earth. More seriously, we are told that from the evidence of their "scat" the giant ground sloths may have become extinct from a drug overdose, through eating the plant known as "Mormon tea". So the book goes on, leaping nimbly from one unlikely topic to another — remote sensing to salt "punch-ups", karst erosion to mass extinctions, until finally reaching the geological glory of field-work in (where else for a muse?) the Elysian Fields.

This is an enjoyable book, a cheerful somersault rather than a solemn symposium and better for entertainment than for education. But that, presumably, was the author's intention. □

Derek Ager is Research Professor in the Department of Geology, University College of Swansea, Swansea SA2 8PP, UK.

Wilder shores

Peter J. Smith

Islands. By H. W. Menard. *W. H. Freeman*: 1986. Pp. 230. £15.95. In the United States available only to members of the *Scientific American Library*.

THERE seem to be as many books called *Islands* as there are islands, concentrating on everything from scenery, through seafaring and survival, to science. Menard's version falls into the last category, as one might expect from a distinguished geological oceanographer who ended up as Director of the US Geological Survey. But a scientific book on islands sounds a bit like one on kidneys; one can see the attraction of specialization whilst taking leave to doubt whether much sense can be made of the parts in isolation from the body as a whole. As far as the Earth sciences are concerned, the revolution of the 1960s, which produced the first convincing global theory of the Earth's behaviour, placed islands firmly in the context of spreading sea-floors and jostling tectonic plates, from which conjunction they are never likely to escape.

Inevitably, therefore, Menard has ended up writing a book largely about plate tectonic processes, albeit with islands taking centre stage and with little mention of continents. Excluded too, incidentally, are those islands, large or small, with a continental structure (Britain, New Zealand, Japan, the Seychelles, certain Indonesian islands and so on). Menard's islands are oceanic — volcanic or volcanic/coral, "small, young, isolated, simple, and subjected to a limited range of environmental factors". In geological terms they represent the closest that nature comes to

laboratory conditions, with dependent variables at a minimum.

Notwithstanding the restricted definition, there is a surprising amount to say. For one thing, islands are perpetually under the influence of spreading lithosphere and stationary hot spots, and all that means in terms of aging, subsidence, rates of extrusion, distribution, redistribution and, ultimately, subduction, continental underplating or terrane formation. New volcanic masses lead to isostatic adjustments and lithospheric warping, the effects of which can be obscured by multi-causal changes in sea level. Then there is the whole phenomenology of underwater volcano formation and decay, involving magma flow, growth rates, erosion and submergence, with or without the participation of coral. Finally, the issue of colonization raises important questions about evolution, migration and civilization.

Plate tectonics though much of this may be in essence, it is plate tectonics viewed obliquely. It does no harm at all to be forced to look at familiar concepts from an unusual angle as long as the viewpoint is coherent, which is Menard's hands it certainly is. The level could best be described as *Scientific American*, which means that the absolute novice may find it tough going in parts. Those with a smattering of basic terminology, however, will discover a rewarding and unusually enjoyable volume, not at all a bad memorial to an Earth scientist who was ever ready to look beyond his immediate research. □

Peter J. Smith is Reader in Earth Sciences at the Open University, Walton Hall, Milton Keynes MK7 6AA, UK.

● Another of Henry Menard's books to be published posthumously, *The Ocean of Truth: A Personal History of Global Tectonics* (Princeton University Press), will be reviewed in a future issue of *Nature*.

Feathered flights of fancy

Tom Kemp

Archaeopteryx, The Primordial Bird: A Case of Fossil Forgery. By Fred Hoyle and Chandra Wickramasinghe. Christopher Davies, PO Box 403, Sketty, Swansea SA2 9BE, UK:1986. Pp.135. £10.95.

THE eminent British astronomers Fred Hoyle and Chandra Wickramasinghe have greatly exercised the vertebrate palaeontology community with their assertion that the famous reptile/bird fossil, *Archaeopteryx*, is a fake. If true, the claim would match the enormity of the fraud associated with Piltdown man. Never having studied any of the specimens of *Archaeopteryx*, nor staked my reputation upon its genuineness, I can cheerfully accept the possibility in principle. Briefly, what is supposed to have happened is that the fossil of a small dinosaur was taken, shallow areas of matrix gouged out and replaced by a paste made from ground-up matrix, and then modern feathers pressed onto the paste to create the imprints now seen.

The authors go about the task of proving their theory with all the comprehensive enthusiasm of fictional detectives, although of the Inspector Clouseau rather than Sherlock Holmes variety. For, whether through perversity or stupidity I do not know, they exhibit a staggering ignorance about the nature of fossils and fossilization processes. Others have commented in detail on the exact misunderstandings (see the paper by A. J. Charig *et al.* in *Science* **232**, 622-626;1986). In general terms, Hoyle and Wickramasinghe draw our attention to such points as changes in the texture and coloration of different parts of the matrix, failure of the part and counterpart slabs to match precisely, the "impossibly" large areas of feather impressions exposed on the slab and the "double impression" of some of the feathers. These are illustrated by the excellent photographs, although not always as convincingly as the text claims. The point is, however, that if features such as these are taken as evidence for forgery, then I do not believe that I have ever seen a genuine vertebrate fossil in my life! They all reflect quite common, perfectly explicable aspects of the process of formation of a fossil.

Two things interest me about this otherwise worthless book: why it was written, and what effect it might have. The motive seems to relate to the authors' own theory of evolution, described very briefly in the first chapter. It seems that major evolutionary change occurs as a result of "genetic storms", whence showers of virus parti-

cles from outer space invade the Earth and become incorporated into the genomes of living organisms. That this theory blithely ignores the multitude of anomalies it creates out of phenomena at present perfectly well explained by conventional theory is set aside, despite the glib invocation of Occam's razor. Apparently, periods of mass extinction as well as major evolutionary change result from large genetic storms, and therefore the origin of a group such as the birds must have coincided with a mass extinction. The nearest relevant-looking mass extinction was at the end of the Cretaceous; therefore there could not have been any birds in the Upper Jurassic; therefore *Archaeopteryx* must be a fake (and the toothed birds of the Cretaceous misidentified to boot). To support this conclusion the most outrageously improbable conspiracy theory has to be invoked. For more than a century palaeontologists seem to have been involved in a Watergate style cover-up; heavy hints that the present staff of the British Museum (Natural History) are still cheating and lying are not excluded. Do astronomers really apply this standard of reasoning in their own field?

As to the effect of the book, I am afraid that it has a superficial appeal and will undoubtedly gain currency amongst Creationists ("Now that we know *Archaeopteryx* to be a fake..."). Thus will science,

the scientific method and the cultural importance of trying to understand our world suffer yet another blow. At best this book is mischievous; at worst it might well be described as evil, because it betrays that morality of objectivity that science cherishes. Perhaps *Archaeopteryx* is a fake. Certainly the possibility should be investigated. But it should be done by those who actually understand fossils, fossilization and fossil preparation, not by a couple of people who exhibit nothing more than a Gargantuan conceit that they are clever enough to solve other people's problems for them, when they do not even begin to recognize the nature and complexity of those problems. □

Tom Kemp is Curator of the University Museum, Parks Road, Oxford OX1 3PW, UK, and a Lecturer in the Department of Zoology, University of Oxford.

New in paperback

- *Biophilia: The Human Bond with Other Species* by Edward O. Wilson (reviewed in *Nature* **312**, 205; 1984). Publisher is Harvard University Press, price is \$7.95, £5.95.
- *Hawks, Doves, and Owls: An Agenda for Avoiding Nuclear War* edited by Graham T. Allison, Albert Carnesale and Joseph S. Nye, Jr (reviewed in *Nature* **318**, 415; 1985). Publisher is W. W. Norton, price is £5.75, \$6.95.
- *Quantum Field Theory* by L. Ryder (reviewed in *Nature* **315**, 610; 1985). Publisher is Cambridge University Press, price is £15, \$24.95.

Ornithology from Princeton

A Guide to the Birds of Colombia

Steven L. Hilty and William L. Brown
Plates by Guy Tudor

H. Wayne Trimm, John Gwynne, Larry McQueen, John Yrizarry, and P. Prall

Line Drawings by Michel Kleinbaum and John Gwynne

Describing all of Colombia's birds, Steven Hilty and William Brown bring together information on one of the world's largest avifaunas — nearly 1,700 species. Over half of all the species of birds in South America are included, thus making the book useful in regions adjacent to Colombia, as well as in the country itself.

836 pages. 56 color plates, 13 halftone plates, 99 line illus., 2 maps. Paperback edition sewn for extra strength.
P: \$42.50. C: \$95.00

At your bookstore or

Princeton University Press 41 William Street Princeton, NJ 08540

Birds of New Guinea

Bruce M. Beehler, Thane K. Pratt, and Dale A. Zimmerman
Illustrations by Dale A. Zimmerman and James Coe

This book is the first definitive ornithological field guide to New Guinea, the world's second largest island and one of the richest tropical environments in the Old World. Because of its isolation, New Guinea's biota includes spectacular radiations of unusual plants and animals (of which the birds of paradise are perhaps best known). This book treats all of the more than 700 species of birds recorded from the region, illustrating more than 600 of them in forty-seven fine color plates and eight black-and-white halftone plates.

293 pages. 47 color plates, 8 halftone plates, 16 line illus., 4 maps plus endpapers. Paperback edition sewn for extra strength.
P: \$37.50. C: 65.00

Putting flesh onto the bones

Malcolm C. McKenna

Mammal Evolution: An Illustrated Guide. By R. J. G. Savage and M. R. Long. *British Museum (Natural History)/Facts on File*:1986. Pp.264. £17.50, \$35.

FOR most fossil mammals, dry bones have been brought to life inadequately by professional palaeontologists. Among extinct animals, the dinosaurs have stolen most of the limelight, and much of the rich store of research-level information about mammalian evolution reaches the public only as background material to the latest discoveries of rival anthropologists. Although there is a great story to tell about mammalian evolution, and a great opportunity to discuss how the scientific method can be applied to the subject, the good books on mammalian evolution over geological time have long been out of print or are badly outdated. Palaeontologists are bombarded by requests about the appearance in life of extinct mammals, and how they changed with time, but we possess few attempted reconstructions to show to the public; those that we do have are generally clones of previous attempts. Although professionals sometimes wince at reconstructions of fur colour patterns and other indefensible extrapolations from what is actually preserved in the fossil record, such images are undeniably useful for non-specialists. Life is breathed into the subject.

In this book, Michael Long (artist) and R. J. G. Savage (palaeontologist) do bring

the evolution of mammals to life, from their poorly known beginnings in the Mesozoic to the richly known record of the Cenozoic. The book combines evocative restorations with spritely text. Many of the illustrations are new and plausible reconstructions of unfamiliar beasts; others clearly owe a debt to previous renderings. Some are outrageously funny, such as the human-faced *Thylacosmilus nicki*, or the contrite *Gigantopithecus* and wistful *Lystrosaurus*. The text is well written and pitched at the right level to maintain interest, mixing solid information with interesting references and speculation. It starts with a quotation from Shelley's *Prometheus Unbound* and ends with Dale Russell's silly illustration of a hypothetical ET-like "dinosauroid hominid". Between these there is much food for thought for non-specialists. There are a number of minor errors, but these can be weeded out for the future editions that will surely follow.

Our understanding of mammalian evolution is itself evolving, as more complete specimens are found and, perhaps more importantly, as methods of analysis improve. Although I found the book rewarding, one criticism is that the authors have missed the chance to move up from Kiplingesque scenarios to epistemologically more mature ground. After all, the British Museum (Natural History) is leading the way in injecting scientific method into its exhibits about evolution. To complement that effort, the book could have moved from exciting animals to greater emphasis on exciting problems. □

Malcolm C. McKenna is Frick Curator in the Department of Vertebrate Paleontology, American Museum of Natural History, New York, New York 10024, USA.

Edges of knowledge

A. Hallam

Controversy in Victorian Geology: The Cambrian-Silurian Dispute. By James A. Secord. *Princeton University Press*:1986. Pp.363. \$49.50, £33.10.

THOUGH geologists for the most part do not exhibit much interest in the history of their subject, there are few, in Britain at least, who do not know a little about the battle of those nineteenth-century giants Adam Sedgwick and Roderick Murchison. What had begun in friendly collaboration gradually developed over a quarter of a century into a bitter feud resulting in mutual estrangement. Like many stories known only in outline, however, the Cambrian-Silurian controversy is surrounded by misconceptions or is not properly appreciated for its most essential characteristic: it was a debate about the proper way to conduct a particular type of science, rather than a mere territorial or priority dispute. With an abundance of letters, notebooks and accounts of meetings available for study, the subject has long been ripe for the sort of book that Secord has written, a detailed, scholarly analysis using modern insights into the sociology of Victorian science.

Through the application of the novel technique of correlating strata by fossils, by the early 1830s great strides had been made in establishing both a relative time-scale and producing a reliable geological map for the structurally uncomplicated younger rocks in Britain. The ancient "grauwacke" or transition rocks of Wales, however, were a geological *terra incognita* that provided a kind of new frontier challenge. Sedgwick started work in North Wales and Murchison in the Welsh Borders, which led respectively to their proposing the Cambrian and Silurian systems, with a mutually agreed boundary being established. After some years it became apparent that Sedgwick's Upper Cambrian corresponded with Murchison's Lower Silurian. The newly created Geological Survey adopted Murchison's classification and Sedgwick became increasingly isolated. But the turning point came with the recognition of a major unconformity between Murchison's Lower and Upper Silurian, and the discovery of "primordial" fossils at a stratigraphic level well below that in which anything had previously been found. In the 1850s and 1860s many leading geologists switched their allegiance to a three-fold classification of what we now call the Lower Palaeozoic, and the dispute effectively came to an end in 1879 when Charles Lapworth proposed the term Ordovician for Murchison's Lower Silurian.

Secord argues that "Cambria" and



Restored to life—
Brontotherium, a
browsing ungulate from
the early Oligocene of
North America. The
animal stood 2.5 m at the
shoulder and was the
rhinoceros of its day.

"Siluria" were conceived of and valued as labels for scientific properties and signals of accomplishment in the quest for originality, with Sedgwick fighting for the Cambrian as a compensation for his failure to complete a major geological book and the more urbane but autocratic Murchison exhibiting a kind of imperialist expansionism in spreading the Silurian both across the world and down to the base of the oldest recognizable fossil-bearing rocks. Whereas Murchison laid emphasis on fossils as the only sound basis for correlation, Sedgwick thought it imperative to establish rock sections with a variety of physical criteria in a way that exemplified the principles of induction that he taught his students at Cambridge; the vehement defence of his Cambrian territory was a kind of moral crusade for the principles of truth and justice as he saw them.

Throughout the dispute, we perceive a characteristic Victorian preoccupation with facts and their classification, rather than with the testing of theories involving process and time. Secord brings out clearly the influence of various social factors in affecting the position which particular people adopted, such as personal com-

petition, the character of the institution to which they belonged, the opportunities for professional advancement and so on. Furthermore, the nature of the underlying assumptions, and indeed the whole conceptual framework, had an influence on what were recognized as "objective" facts.

Secord adopts a straightforward narrative approach and has produced a well-written, sometimes absorbing account which demonstrates convincingly the heuristic value of controversy in promoting scientific advance. What I missed, however, was more of the authentic voice of the leading disputants. Sedgwick was by all accounts a lively and engaging, if somewhat choleric, person who expressed himself with great force and eloquence. It would have been gratifying to see more evidence of this through the inclusion of extended extracts from his correspondence and notebooks, and to contrast it with the more restrained but dogmatic tone of Murchison who is likewise poorly served by quotation. □

A. Hallam is Lapworth Professor of Geology, University of Birmingham, PO Box 363, Birmingham B15 2ET, UK.

Top people read...

John C. Marshall

The Harvard Guide to Influential Books: 113 Eminent Harvard Professors Discuss The Books That Have Shaped Their Thinking. Edited by C. Maury Devine, Claudia M. Dissel and Kim D. Parrish. Harper & Row: 1986. Pp. 300. Hbk \$18.95; pbk \$7.95.

WHAT do *Rats, Lice, and History*, *Dynamic Administration*, *Amateur Telescope Making* and the *Adventures of Pinocchio* have in common? The answer is that each was chosen by one of "113 eminent Harvard professors", who here "discuss the books that have shaped their thinking". I emphasize "one" because the most striking characteristic of *The Harvard Guide to Influential Books* is how few choices the contributors hold in common. Between them, the distinguished professors refer to some 560 works, most of which I (and you) have not read. I accordingly thought that the best way to educate myself would be to pick out the *really* influential books, say those that at least five contributors mention. We are now left with *The King James Bible* (seven listings), *War and Peace* (six) and *The Republic*, *The Federalist Papers* and *The Adventures of Huckleberry Finn* (five each). If authors with listings for multiple books are counted, Winston Churchill, Sigmund Freud, Thomas Mann, William Shakespeare and Max Weber just manage to creep in.

To catch a scientist we must drop to four citations for *On the Origin of Species*. Physicists and chemists are conspicuous mainly by their absence. No one seems to have been much impressed by Galileo, Newton or Max Planck, nor inspired by Lavoisier, Dalton or Mendeleev. Gone are the days when eminent Harvard mathematician Tom Lehrer concluded his song of the elements with the immortal lines: "These are the only ones of which the news has come to Harvard. There may be many others but they haven't been dicarvard". Einstein is cited once, by Nicolaas Bloembergen, for his 1938 book (with Infeld) *Physik als Abenteuer der Erkenntnis*. Other physicists enter either as philosophers of science or as (proto-) molecular biologists. (No prizes for guessing Schrödinger's *What is Life?* in the latter category.) Mathematicians figure only as statisticians and "decision theorists" (*vide* Von Neumann and Morgenstern's *Theory of Games and Economic Behavior*). No one's life was made more (or less) uncertain by Heisenberg's principle; no one was converted to existentialism by Gödel's theorem.

In part, the obvious omissions result from the editors' admonition that contributors should not merely recite "Western Civilization's Great Books", but rather tell of the printed matter that *actually* influenced them. Much of the compilation's charm arises from this constraint. Thus architect Moshe Safdie engagingly praises Rudofsky's *Architecture without Architects*; E.O. Wilson confesses that, at age sixteen, Lysenko's *Heredity and Its Varia-*

bility enchanted him with its "radical and effective challenge to conventional science"; and, best of all, Mary Chatfield cites "any good cookbook" as evidence that "the rational is still achievable in the most chaotic of worlds". Occasionally one is perplexed by the reasons given for a particular choice. Historian Bernard Bailyn admires Ernest Jones's *Life and Work of Sigmund Freud* as "a morality tale of... heroic achievements", but spoils the effect by concluding "And it happens to be true". For psychologist Richard Herrnstein, *The Trial and The Castle* "captured the lunacy and futility of individuals struggling with bureaucracies". Poor Frank: when will people learn that Kafka is not kafkaesque?

None of the faculty has been shaped by the *Talmud*, the *Life of the Buddha* or the *Koran*; indeed, insofar as the "East" (either near or far) exists at all for the contributors it is usually in the form of economic competition with Japan (*vide*, for example, *Japan as Number One: Lessons for America*). Piety comes in predominantly paternal form, although Diana Eck (a professor of comparative religion) writes thoughtfully about Mary Daly's *Beyond God the Father: Toward a Philosophy of Women's Liberation*, and Sidney Verba (director of the University Library) recommends Chaim Grade's memoirs of the *shtetl* in *Rabbis and Wives*. Philosophy tends to the conservative; when Aristotle, Hayek, Hegel, Mill and Popper turn up it is more for their politics than their metaphysics. The spirit of the Law School, the Business School and the Kennedy School of Government dominates the enterprise. *Das Kapital* receives a derisory two votes, *The General Theory of Employment, Interest, and Money* scores zero. Proudhon, Bakunin and Prince Kropotkin fail to make the grade, whilst *Explaining the Cuban Missile Crisis*, *America's Competitive Edge* and Machiavelli's *Prince* get straight A's. Balance is only marginally restored by educationalist Francis Keppel, who cites *The New York Times*, a choice that serves to remind us that Britain has not had a decent daily since the *Manchester Guardian* dropped the *Manchester*.

Reviewers often decline to take compilations such as this seriously. *Au contraire*, I found the *Harvard Guide* fascinating and deeply informative. My only regret is that space was not found for the book-lists of eminent Harvard secretaries — the people who really keep an institution moving. I eagerly anticipate a series of such volumes from major universities. They would provide a far better guide to advising one's children on which centres of higher learning to apply to than do the usual bland prospectuses. □

John C. Marshall is in the Neuropsychology Unit, part of the Neuroscience Group at the Radcliffe Infirmary, Oxford OX2 6HE, UK.



INSTITUT FRANÇAIS DE RECHERCHE SCIENTIFIQUE
POUR LE DÉVELOPPEMENT EN COOPÉRATION

Guide des étoiles de mer, oursins et autres échinodermes du lagon de Nouvelle-Calédonie. Handbook of the sea-stars, sea-urchins and related echinoderms of New-Caledonia lagoon. A. GUILLE, P. LABOUTE, J.-L. MENOUE. Partly bilingual: French, English. 1985. 240 p., 224 photos coul. 350F.

Cloffa 1. Check-list of the freshwater fishes of Africa. 1984. 410 p. Musée royal de l'Afrique centrale/ORSTOM. 130F.

Cloffa 2. Check-list of the freshwater fishes of Africa. 1986. 520 p. Musée royal de l'Afrique centrale/ORSTOM. 140F.

Mangroves du Sénégal et de la Gambie. Ecologie - Pédologie - Géochimie - Mise en valeur et aménagement. C. MARIUS. 1985. 368 p., 10 pl. photogr. 135F.

Changements globaux en Afrique durant le Quaternaire. Passé - Présent - Futur. INQUA - ASEQUA. Symposium international. Dakar 21-28 avril 1986. 1986. 528 p., some articles in English. 300F.

Monographie hydrologique du fleuve Niger. Ministère des Relations Extérieures. Secrétariat d'Etat à la Coopération. ORSTOM. **Tome I - Niger supérieur.** 1986. 396 p., 4 cart. ann. noir, 7 microfiches.

Tome II - Cuvette lacustre et Niger moyen. 1986. 506 p., 6 cart. ann. noir. (the 2 volumes: 350F).

Pharmacopées traditionnelles en Guyane. Créoles, Palikur, Wayapi. P. GRENAND, C. MORETTI, H. JACQUEMIN. 496 p., 156 photogr. coul., 49 traits. (forthcoming)

EDITIONS DE L'ORSTOM
70 - 74 Route d'Aulnay
93140 BONDY

Daniel by himself

Jacquetta Hawkes

Some Small Harvest: The Memoirs of Glyn Daniel. Thames & Hudson: 1986. Pp. 448. £12.95.

GLYN Daniel, with characteristic sang-froid, has chosen the title of *Some Small Harvest* for one of the bulkiest autobiographies of recent times: 448 pages and over 50 illustrations. The first photograph shows a captivating little boy with bright, enquiring eyes, perhaps capable of impudence, and stoutly booted for the successful passage through worldly life that is to be chronicled in such loving detail.

Here I must declare an interest. Glyn Daniel went up to St John's College, Cambridge, during that dawn of modern archaeology, the early 1930s, in the very year that I had graduated. He came from a small village in South Wales, I from a home within half a mile of St John's. We both read for the new archaeological tripos and were to pursue the subject, he with infinitely greater diligence, for the rest of our lives. So it is that I share his memories of very many of the individuals and the events great and small which fill his book. It is hard for me to judge whether others, who must read without the glow of nostalgia, will find it as appealing as I do.

Glyn Daniel not only has an extraordinarily retentive memory but through 40 years has kept a journal which arms him with a record that is "contemporary and explicit". What is striking about the resulting life story is its range and variety. It is an exceptional man who, emerging from an unworldly yeoman background, could become a Wing Commander in India (as head of Central Photographic Interpretation), a brilliantly successful editor of *Antiquity* and of a long series of books of the best *haute vulgarisation*, an adventurous Steward of St John's, a writer on food and drink and of detective fiction — and then, when many felt that all this might have weakened his scholarly repute, carry off the much-honoured Disney Professorship of Archaeology at Cambridge.

Daniel writes of these achievements, interspersed with amusing episodes and anecdotes, as though he took life easy as the grass grows on the weirs, whereas in reality there was within him an iron determination to succeed and a rare capacity for hard work. He has such an enthusiastic interest in his own doings that he perhaps sometimes falls into triviality and self-indulgence; there are, for instance, too many recipes and wine lists, and accounts of lectures and broadcasts — yet for much of the way he carries one along on a tide of his own enjoyment.

For readers of *Nature*, most attention

should go to Daniel's archaeological researches and judgements over half a century. Here there is some disappointment. He has published several histories of the subject and can be excused from all such generalities. Yet it would, I think, have given more unity to the autobiography if he could have interwoven the development of archaeological knowledge, method and thought with his personal story; in particular, if he could have gone more fully into the new radiocarbon dating which appears to have nullified much of his extensive research into megalithic



Glyn Daniel, aged five — "stoutly booted for the successful passage through worldly life".

monuments, and also examined the present anti-diffusionist fashion. For fashion I believe it to be: how many publications have I read of late where the author is plainly forcing his conclusions into the "local development" mould. It should be added, however, that Glyn Daniel is consistently opposed to what he chooses to call "bullshit" archaeology, the outpourings of those who are honestly deluded or who deliberately indulge in wild fantasies in order to sell their books by the million.

Standing back to judge *Some Small Harvest* as dispassionately as I can, I am convinced that in the long run it will remain valuable as a sociological source-book, a record of Cambridge and especially college life, of travel and conditions in many lands, of the flavour of twentieth-century archaeology — and a highly personal *Who's Who* of scores of the notable characters the author has known. Viewed in this light the very trivia I have criticized will be of unique interest for a future generation, and may be more than forgiven.

Jacquetta Hawkes, Littlecote, Leysbourne, Chipping Camden, Gloucestershire, GL55 6HL, UK, is an archaeologist and author of, among other books, *Mortimer Wheeler: Adventurer in Archaeology* (Weidenfeld & Nicholson, 1982).

Occupied territory

Randall White

The Palaeolithic Settlement of Europe. By Clive Gamble. *Cambridge University Press*: 1986. Pp. 469. Hbk £40, \$65; pbk £15, \$24.95.

ACCORDING to current evidence, Europe was first occupied by hominids 1.5 million years after the first Palaeolithic tool-makers emerged in Africa. There is much that remains to be known about this successful expansion of our species into northern latitudes, but there is also much information presently available. Research on the Palaeolithic settlement of Europe (roughly 750,000 – 12,000 years ago) is published in so many languages, and is based upon such a variety of theories, that it is difficult for any scholar to have a clear overall view of the subject. Clive Gamble's drawing together of knowledge and references from such an intellectual mosaic is in itself a great achievement.

It must be said, however, that the framework on which the book is organized is puzzling, if not idiosyncratic. The first chapter, which begins with a history of European Palaeolithic studies, concludes with a strong advocacy of Binfordian methodological premises. While Binford has had a very positive influence on European Palaeolithic studies, it is a little surprising to find his views uncritically repeated virtually word for word.

The following chapter, "Hunter-Gatherer Regional Systems", takes a detailed look at regional and inter-regional adaptations among contemporary hunter-gatherers. At the end of it, Gamble follows Robert Foley in proposing rejection of the site as an analytical unit for the study of prehistoric hunter-gatherers. This has become a popular position recently; but it deserves much closer scrutiny in light of the fact that sites do correspond to palimpsests of past living communities, a concept critical to regional studies of living peoples. I found this chapter especially troublesome in its heavy emphasis on 1960s ecological thinking. The result is a form of ecological functionalism which views the social world of hunter-gatherers exclusively with reference to the imputed functions of "alliance" and "information strategies". Anthropological theory in the 1980s is not what it was in the 1960s, and for Gamble to cite Marshal Sahlins's ideas of 1968, without acknowledging Sahlins's refutation of the ecological paradigm, betrays an unduly selective approach.

Chapter 3 is a complex presentation of data on environments and resources. Gamble breaks Europe down into nine sub-regions, upon which are based the inter-regional comparisons in the rest of

the book. Discussion of glacial chronology is detailed and excellent. Also here is a listing of the key anatomical and behavioural attributes of the main faunal species, which is very helpful but unfortunately omits certain attributes of importance to human beings, such as hide and antler quality and aggregation behaviour.

In the next, very short chapter, entitled "Technological and Typological Approaches", Gamble explores technological and typological evolution during the European Palaeolithic. The result is a severe revision of the traditional sequence in which the present Middle and Lower Palaeolithic are melded to form the "Early Palaeolithic", which comprises all industries before 35,000 years ago. The terms Lower and Middle Palaeolithic are inadequate, and Gamble's lumping of the periods together clears the decks for future analyses of much potentially interesting variation. It is all the more curious, then, to see him become a splitter when he turns his attention to the Upper Palaeolithic, creating two divisions (before and after 20,000 years BP). The later division is based upon the emerging importance of microliths after 20,000 years BP. Unfortunately, it ignores recent excavations in which Aurignacian (camina-de scrapers) and Gravettian (microgravettes, backed bladelets) levels have yielded abundant microliths.

Chapter 5 is a useful, well-referenced discussion of the data base for research on Palaeolithic Europe, organized according to Gamble's newly developed temporal and geographical structure. Each region is summarized according to (1) chronology, (2) industrial groupings, (3) geographical variation and regional case-studies, and (4) burials and art (though the last of these is rather sketchy).

The sixth chapter, "Space and Subsistence", is the best in the book and probably the most comprehensive account of Palaeolithic settlement and subsistence patterns that has appeared to date. There are two shortcomings, however. First, because of its supposed circularity site-location research is summarily dismissed. Second, after having illustrated some brilliant work on intra-site spatial patterns, Gamble re-states his view that the site should not be our focus of analysis. My position is that the examination of individual sites must be used to build a regional picture of human landscape use.

"Demography and Style", the next chapter, is puzzling, because the section on demography deals almost exclusively with carnivore behaviour as a basis for research into the distribution of populations in space. The section on style and interaction relies heavily on information approaches to variability in material culture, exemplified by Gamble's own work on Gravettian female figurines. Again, in

true 1960s fashion adaptive significance is attributed to form in portable and cave art; based upon a very narrow reading of the Australian literature, the latter is treated as territorial marking rather than as evidence for complex systems of meaning that help to define the very nature of human adaptation.

Chapter 8, "Society, Sediments and Settlement", emphasizes the interesting regional differences in stone-tool and site densities. Here, Gamble wishes to chip away at the notion that a universal set of Upper Palaeolithic traits, rather than local environmental exigencies, conditioned the archaeological record in all regions. This is indeed an important issue, but Gamble's density calculations need to be examined closely before being accepted. For example, his find-spot density for the Perigord is based on a surface area of 12,600 square kilometres, approximately twice the true figure. Moreover, historical differences in the intensity of site search need to be taken into consideration. It is a pity that Gamble fails to suggest the need for systematic controlled survey to control sample quality in assessing overall population increase across the transition between the Middle and Upper Palaeolithic.

In the concluding chapter, the virtues of viewing culture as an adaptive process are extolled, especially as applied to the emergence and diversification of culturally modern humans at about 35,000 years BP. Gamble seems not to recognize that those of us who view the Upper Palaeolithic as an evolutionary threshold are not anti-adaptive; rather, we are comfortable in switching back and forth between discussion of the evolution of culture and the evolution of cultures.

More disturbing in the conclusion, however, are the peculiarities that emerge in the form of three models for what human beings did when they reached Europe. The problem posed by cold northern climates is viewed primarily as a problem of finding a niche, which ignores clothing, fire and shelter as cultural evolutionary prerequisites. In any event, Gamble proceeds to propose that the earliest Europeans were adapted to the consumption of frozen animal carcasses resulting from natural deaths in autumn and winter. He seriously suggests that the wooden "spears" from Leheringen and Clacton were probes for finding dead animals beneath the snow. This is an unfortunate excursion onto the fringe.

While readers must work around some awkward theoretical and organizational impediments, the rich (though already somewhat dated) documentation and first-class presentation of this book will make it a basic reference in English-language Palaeolithic studies for years to come. □

Randall White is an Assistant Professor in the Department of Anthropology, New York University, New York, New York 10003, USA.

Principled man in unprincipled times

Martin J. Klein

The Dilemmas of an Upright Man: Max Planck as Spokesman for German Science.
By John L. Heilbron. University of California Press: 1986. Pp.238. \$16.95, £14.50.

MAX PLANCK (1858–1947) exemplified the traditional values of the world in which he grew up, the world of a conservative, cultured German family of pastors, scholars and jurists. The very model of a classical physicist, he “embodied with rare completeness”, according to Albert Einstein, the “ideal search for truth . . . a bond forever uniting scientists of all times and in all places”. But his very devotion to the search for absolute scientific truth led Planck inexorably to the quantum of energy with which his name is permanently linked, the symbol of this century’s non-classical physics. His unshakable sense of responsibility led him, equally inexorably, to accept a series of positions of authority in the German academic and scientific establishments. By 1914 he had become the acknowledged spokesman for German physics, a role he could not in conscience relinquish when his country was transformed into the Weimar republic and then entered the dark times that began in 1933. As John Heilbron writes in his preface, Planck’s

struggle to come to grips with changing political circumstances, and his efforts to adjust the image of science and civic life that he developed before 1900 to the realities of twentieth-century Germany have the elements of heroic tragedy.

Heilbron’s book, the first in English devoted to Planck, is an outstanding piece of work by one of our finest historians of science. Heilbron’s earlier writings include an admirable study of Henry Moseley’s life and work, a major book on early intellectual and institutional developments in physics, *Electricity in the 17th and 18th Centuries*, and a series of articles and monographs on a wide range of subjects in the physical science of the past four centuries. Among its other impressive features, his work always shows his mastery of “the bibliographic resourcefulness proper to the historian”.

As his title implies, Heilbron’s primary concern in this book is Planck’s role as spokesman for science rather than the physics he actually did. Heilbron does sketch Planck’s development from exclusive devotee of thermodynamics through his gradual acceptance of the need for both atomism and a statistical

approach, and he shows how Planck used his hard-won theory of blackbody radiation to determine the first reliable values of the atomic constants. We see Planck, conservative in his stance towards the developing quantum theory, immediately and enthusiastically respond to the special theory of relativity and its

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Max Planck — public silence, private integrity.

author. Planck was one of those who brought Einstein to Berlin in 1914, and it was Planck’s presence that helped to keep him there: “To be near Planck is a joy”, Einstein once wrote.

There is no irony in Heilbron’s designation of Planck as an upright man. “Respect for laws, trust in established institutions, observance of duty, and absolute honesty” are listed as “the hallmarks of Planck’s character”. These were combined with genuine modesty and the clear conscience that came with his responsible adherence to his principles. Not a man for public pronouncements, Planck showed his opposition to academic anti-Semitism preceding the First World War by acting quietly to keep qualified Jewish physicists such as Emil Warburg

and Lise Meitner from being denied positions on irrelevant grounds.

By 1914 Planck held a series of major positions in the German Physical Society, the Prussian Academy of Sciences, the Kaiser-Wilhelm-Gesellschaft, and, as war began, as Rector of the University of Berlin. He faced the war with mixed emotions, foreseeing “something great, along with something monstrous”. Planck was one of the 93 intellectuals who signed the famous “Appeal to the Cultured Peoples of the World” in October 1914, denying German atrocities in Belgium and expressing the unity of German culture and German arms. Unlike most of his co-signers he soon moderated his position and worked to maintain the possibility of post-war international reconciliation among scientists. His 1916 open letter to H.A. Lorentz emphasized his

firm conviction that there are domains of intellectual and moral life that lie beyond the struggles of nations, and that honourable co-operation in the cultivation of these international cultural values and, not less, personal respect for citizens of enemy states are indeed compatible with ardent love and energetic work for one’s own country.

Few citizens of any belligerent nation have publicly subscribed to such enlightened views in the midst of a modern war.

Planck had to call on all his internal resources to go on with his work after his own losses in the war years: his first son was killed in action and his twin daughters both died in childbirth. Devotion to duty sustained him in this time of personal grief and national upheaval. Unlike many of his conservative colleagues he gave his loyal support to a republic whose principles were not his own. Heilbron discusses Planck’s staunch support of Einstein when vicious anti-Semitic attacks were directed against both the man and his general relativity, but he stresses Planck’s avoidance of public statements that he thought might encourage the politicization of science. We see Planck struggling to help restore normal international relations among scientists despite the severe restrictions imposed by the victorious Allied powers. As if these tasks and his efforts to secure adequate financial support for physicists were not enough for a man approaching his seventieth year, Planck also undertook to defend a causal, realistic philosophy of physics in the new age of quantum mechanics.

The problems Planck dealt with in the Weimar era were minor compared to those he confronted after Hitler took

Mary Evans

power in 1933. As secretary of the Academy and president of the Kaiser-Wilhelm-Gesellschaft, Planck could not avoid dealing with the government. Convinced at first that the Nazi excesses could not last, and that Hitler's call for national revival and unity had much to be said in its favour, Planck was ready to accommodate to the demands of the new regime. At 75 he would have been relieved to give up the burdens of his offices, but he felt responsible to those who counted on his help and trusted his "immensely noble purity".

Planck's hopes for quick return to civilized behaviour in Germany were quickly dashed when his Jewish colleagues were forced out of their positions and many of them left the country. Still reluctant to make public statements, Planck did not protest against the Academy's offensive response to Einstein's resignation. He went on advising his younger colleagues to have patience and stay at their posts. Nevertheless, it was Planck who went directly to Hitler in May 1933 and tried to convince him that his policies were destroying German science. Hitler would make no distinctions, and became so enraged that Planck could only "fall silent and take my leave". His efforts to salvage what could be salvaged, to block (behind the scenes) the irrational behaviour of his fellow Nobel laureates Philipp Lenard and Johannes Stark, both enthusiastic admirers of Hitler since 1924, did have some positive results. More and more frequently, however, he had to compromise his principles and accept the unacceptable.

During the Second World War, now an octogenarian but free of offices at last, Planck continued his public lectures on science and philosophy, using them as a platform for defending the values of science and civilization and for implicit criticisms of the regime. The disasters of the war, and especially the arrest and execution of his son Erwin for complicity in the 1944 attempt to assassinate Hitler, finally broke Planck's spirit. He died in 1947, just short of his ninetieth birthday.

Heilbron's thoughtful book makes it clear that there were no easy answers to the dilemmas Planck faced. Einstein did not forgive Planck for his public silence in 1933, but he had never lived with Planck's responsibilities. One can speculate on what might have happened if Planck had made a different series of decisions, but speculation is only speculation. Planck acted in accordance with his principles and his conscience, though the world crumbling around him had little place for either. Heilbron writes in his preface that Planck's "life is a lesson", but each reader must decide for himself what that lesson is. □

Martin J. Klein is Higgins Professor of the History of Physics at Yale University, Box 2036, New Haven, Connecticut 06520, USA.

Matter of utmost gravity

Joseph H. Taylor

Was Einstein Right? Putting General Relativity to the Test. By Clifford M. Will. Basic Books: 1986. Pp.253. \$18.95.

IN THE centenary year of Albert Einstein's birth, 1979, a number of books appeared which described the man and his science, especially general relativity. Seven years later, Clifford Will has added another to the list — a readable volume detailing the renaissance that has taken place in gravitational physics since a few years after Einstein's death in 1955. Writing for a general audience, Will describes (sometimes with first-hand knowledge, and always with flair) developments of the past quarter-century which provide the underpinnings of our understanding of the gravitational interaction in nature and its apparently inextricable connection with the geometry of the Universe.

The 1960s, heyday of radio astronomy, dawn of the space age and beginning of the era of atomic clocks, saw for the first time the development of experimental techniques well suited to the high-precision measurement of tiny differences between the predictions of newtonian gravitation theory, general relativity and other "currently viable alternatives". Will tells us of these developments, taking a topical approach. He starts with the reincarnated "Eötvös experiments" carried out in Princeton and the Soviet Union, which verified the equivalence principle to very high precision. Gravitational red-shift tests are described next, with emphasis on the Pound-Rebka-Snider experiments using gamma rays in the Jefferson Tower at Harvard, and the "flying atomic clocks" experiments carried out by Robert Vessot and others a few years later. Together, these tests showed conclusively that a correct theory of gravity must be a metric theory, embodying the concepts of curved space-time.

Subsequent chapters deal with modern experiments measuring the bending of radio waves by the Sun's gravity, the perihelion shift of Mercury by radar ranging techniques, the slowed propagation of signals in the solar gravitational field, extension of the equivalence principle to include the self-gravitational binding energies of the Earth and Moon, the constancy of the gravitational "constant" G , and the loss of orbital energy in the binary pulsar system by gravitational radiation. Throughout the narrative, Will sprinkles a trail of personalities and anecdotes. Robert Dicke's name appears frequently: first as the author of a 1961 article in *Scientific American* which helped to convert the

teenaged Clifford Will from a career in architecture, or maybe genetics, to one in physics; later as quintessential experimenter, and co-author of a theory of gravitation to rival Einstein's; later still as a passenger on an airplane bound for Missoula, Montana, on which he was informed by Kenneth Nordvedt, then a total stranger, that the Brans-Dicke theory appeared to be inconsistent with the equivalence principle for strongly self-gravitating bodies.

Irwin Shapiro's discovery of the gravitational propagation delay for electromagnetic waves makes a good story; it includes an image of Shapiro making an eleventh-hour pitch — via long-distance telephone connection from Cambridge,



Radar ranger — the Millstone Hill antenna, northern Massachusetts, used by researchers at the Massachusetts Institute of Technology from 1959 for radar echo observations of the planets.

Massachusetts, to the Jet Propulsion Laboratory in California, complete with locally shown viewgraphs, NASA style — pleading for inclusion of the time-delay experiment on the forthcoming Viking mission to Mars. The story of the discovery of the Hulse-Taylor binary pulsar is here, too, complete with graphical sketches reproduced from Russell Hulse's laboratory notebook and with the correct inference that graduate student Hulse and assistant professor Taylor were thoroughly astonished by the potential implications of what they had found.

Will has succeeded in what he set out to do, namely to write an interesting and accessible account of recent progress in one field of physics. □

Joseph H. Taylor is James S. McDonnell Distinguished University Professor of Physics at Princeton University, Princeton, New Jersey 08544, USA.

EXPLORING ENERGY . . .

LEARNING ABOUT ENERGY

by David J. Rose

A wide-ranging text for senior and graduate students that weaves technical and scientific information into an integrated current account of the field, including energy in its social and economic contexts, environmental consequences of energy use, energy conservation, fossil fuel resources and reserves, coal, nuclear power, solar power, and global electrification. A volume in the series *Modern Perspectives in Energy*.

0-306-42124-0/\$59.50
(\$71.40 outside US & Canada)
text adoption price on orders of six or more copies: \$39.50

NEW FROM PLENUM TRADE

EINSTEIN'S DREAM

The Search for a Unified Theory of the Universe

by Barry Parker

"Completely non-mathematical, and beautifully written. Every page is a pleasure."

— Isaac Asimov

Award-winning writer Barry Parker depicts the men and their theories in the search for ultimate answers in modern physics.

0-306-42343-X/\$18.95
(\$22.73 outside US & Canada)

SCIENCE TRIVIA

From Anteaters to Zeppelins

by Charles J. Cazeau

"A work of reference and a delightful bedside book."

— Glyn Daniel, *University of Cambridge*

Let's Explore columnist to 81 Gannett newspapers presents questions that will pique the curiosities of the Ph.D. as well as the grade-schooler.

0-306-42353-7/\$17.95
(\$21.54 outside US & Canada)



Plenum Publishing Corporation
233 Spring Street
New York, N.Y. 10013-1578

In the United Kingdom:
88/90 Middlesex Street
London E1 7EZ, England

Doctor, doctor...

Nicholas A. Rupke

Frontiers of Learning: The Ph.D. Octopus. By Pjotr Hesseling. *Foris Publications*. * 1986. Pp. 204. Hbk £27.95, \$44; pbk £17.95, \$27.

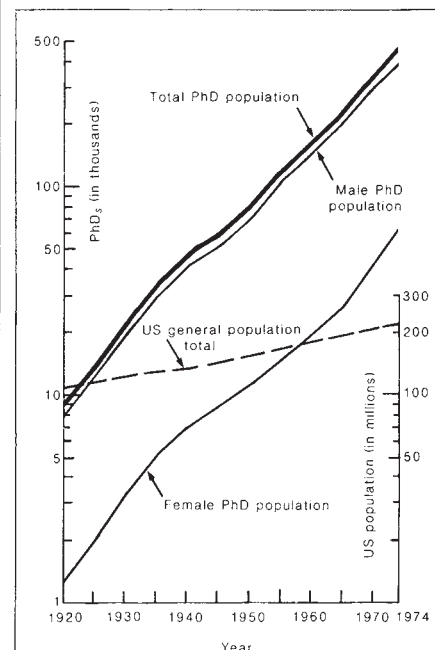
"CANDIDATES must have a PhD . . .". This stipulation has become a standard phrase in the job advertisement columns of science magazines. The PhD, it would appear, has come to represent something of a passport to academic jobs and a union card for the would-be industrial scientist. Employers routinely use the qualification as a preliminary screening device to whittle down large numbers of applicants, often for jobs in which research training has little relevance. Although a result has been a steep rise in the number of people holding a doctorate, especially in the United States, relatively little is known about the PhD phenomenon. What is the function of the degree, its intrinsic worth, as the crowning feature of higher education? What fruits do we expect it to yield? What precisely should it represent? Is the recent multiplication of doctoral degrees a sign of academic health or a craze that is devaluing them?

It is questions such as these that Pjotr Hesseling, professor of business economics at the Erasmus University of Rotterdam, addresses in this book. He uses a comparative approach, looking at the history of the PhD in different countries and cultures. In its modern form the PhD is a German invention, the foundation of the University of Berlin in 1810 marking the beginning of higher education centred on the research imperative. The doctorate became less of an ornamental initiation rite and more of a serious research accomplishment, representing an independent and original contribution to knowledge, embodied in a book-length dissertation. From Germany the PhD moved to the United States (first award in 1861), Britain (1917) and more recently to Japan (1953), gradually penetrating the higher education system of each country.

In the course of its global advance, the PhD has been adapted to various indigenous cultural traditions. The European doctorate is to some extent still a shibboleth of originality and cultural authority (although the precise appreciation varies between disciplines and countries); the American doctorate is primarily used for pre-career training and as a certificate for entering the academic and other professions; the Asian doctorate, which Hesseling has studied mainly in Indonesia and Japan, is still in a formative phase, being

made to fit old traditions of scholarship and wisdom on the one hand, and new, pragmatic considerations of state-managed science policy on the other. Despite these differences, the American graduate school, envied for its productive efficiency, is increasingly becoming the model for the rest of the world.

Hesseling regrets the "Americanization" of the PhD, which is turning it into market-orientated training during which a candidate will often be assigned a narrowly defined topic within a pre-existing research programme. The PhD, he believes,



PhD explosion — the rise in the number of living holders of doctorates earned in the United States between 1920 and 1974.

is becoming too much part of government-directed science policy and of university administration preoccupied with production costs. Quality should be what matters most, but this feature is too elusive to be regulated to any great extent. Originality in research, although the outcome will be less certain, should have priority over completion rates — the reverse of what a recent OECD report, *The Role and Function of Universities: Post-graduate Education in the 1980s*, says on the subject.

Hesseling does not believe that the doctoral dissertation is a necessary complement to the education of all post-graduates. To him, the PhD should be part of a special "cultural mission" which includes testing received knowledge, overcoming boundaries between disciplines, institutions and nations, and striving for such ideals as disinterestedness and communalism. A doctorate should explore, as the title of his book states, "The Frontiers of Learning" and not be an overspecialized tentacle of "The Ph.D. Octopus" (a sub-title borrowed from the Harvard psychologist William James).

*PO Box 509, 3300 AM Dordrecht, The Netherlands, and PO Box 5904, Providence, Rhode Island 02903, USA.

To achieve this end, Hesseling advocates a European doctorate through the revitalization of the traditional academic peregrination which made use of the advantages of different systems. A code of practice should be adopted by the main European universities with, among other things, the abolition of fixed completion periods for doctoral projects, emphasis on interdisciplinary work, freedom for the candidate to select both topic and supervisors (the latter from more than one university), use of external examiners (already the practice in Britain) and publication of the thesis in a major language.

One can readily sympathize with these ideals. But given the increasingly tight and competitive marketplace, and given also the domination of the American system, the need to hold a doctorate as well as a bachelor's and a master's degree is greater than ever. It would therefore seem unrealistic to revert to a system of higher education in which the PhD is for the relatively few. Furthermore, Hesseling's ideals are to a large extent accommodated, especially in the Anglo-American

world, by the network of post-doctoral fellowships which provide an opportunity to elaborate and publish doctoral theses. Many of these fellowships are held in institutes for advanced study with their interdisciplinary and international atmosphere. The Princeton Institute has been replicated, *mutatis mutandis*, not only in the United States itself but also in Germany, The Netherlands and Scotland. In England, several of the Oxbridge colleges give hospitality to advanced researchers in different disciplines.

Closer examination of these and related questions is obviously needed. Hesseling's book is seriously flawed stylistically — there are innumerable mistakes of English usage and unnecessary repetitions of material, and many passages are exasperatingly vague. But he has brought a commendable breadth of approach to his subject, and has provided much food for thought for anyone interested in higher education. □

Nicholas A. Rupke, Wolfson College, Oxford, OX2 6UD, UK, is a geologist and historian of science and medicine.

Quantum question

Philip Pearle

The Shaky Game: Einstein, Realism, and the Quantum Theory. By Arthur Fine. University of Chicago Press: 1986. Pp.179. \$25, £21.25.

EVEN such a loving biographer of Einstein and his physics as Abraham Pais seems embarrassed by Einstein's position on quantum theory. In his book *'Subtle is the Lord'*, Pais quotes Einstein — "I feel that the real joke that the eternal inventor of enigmas has presented us with has absolutely not been understood as yet" — and then remarks, "It is believed by nearly all of us that the joke was understood soon after 1925 . . .". Given this belief, and the spectacular successes of quantum theory which nurture it, it is perhaps not surprising that sympathetic and careful analyses of Einstein's attitude have been lacking.

This lack is quite satisfactorily addressed in five (the second to sixth) of the nine essays which make up this book. The essays, most of which have been previously published, are clear, well-reasoned and appropriate settings for the gems they are studded with, quotes from the Einstein archives at Princeton. The sixth essay, "Einstein's Realism", forms a bridge to a rather different piece of country, three provocative articles on realism which reject both it and anti-realism in favour of the author's own position, which he calls the "Natural Ontological Attitude" (NOA).

The second essay, which follows an

introductory chapter, attacks the often-promulgated view that Einstein's position on quantum theory represented a "kind of scientific senility". After giving ample evidence that Einstein's way of thinking supported Poincaré's admiring assessment of ". . . the facility with which he adapts himself to new concepts. . . . He has not remained attached to classical concepts . . .", Fine refreshingly concludes that ". . . it was Bohr not Einstein who felt bound by the classical concepts", and "in the end Einstein was more radical in his thinking . . .".

The third essay details Einstein's route to the conviction that quantum theory is not ". . . in principle capable of producing a complete description of an individual physical system", and culminates in a discussion of the famous Einstein-Podolsky-Rosen paper. For those of us who have secretly never felt fully at ease with the tortuous arguments in that paper, and who have wondered how that master of clarity, Einstein, could have done this to us, Fine brings welcome relief. He shows just how complicated the argument of the paper is, presents evidence that Podolsky wrote it without revision by Einstein and analyses a simpler argument that Einstein sent to Schrödinger in a letter. The fifth essay also concerns letters from Einstein which stimulated Schrödinger while he was writing his famous paper on the cat paradox. This paper showed that quantum theory can predict a state for a cat that is a superposition of cat-alive plus cat-dead, an unrealistic situation for a single cat, thus supporting Einstein's statistical interpretation of quantum theory: ". . . the description not of a single system but of an

ensemble of systems". How delightful to learn that Einstein himself initiated this paradox, with an example of unstable gunpowder in a quantum state of unexploded plus exploded!

I found essay four on Einstein's statistical interpretation to be rather incomplete. Perhaps because Einstein felt that the de Broglie-Bohm "hidden variable" model is not fundamental enough, no discussion is given of this or the models of Wiener-Siegel or Nelson, which many feel are nice examples of how a statistical interpretation may function. Instead Fine presents some of his own work, a refutation of a model he calls "The Standard Answer" (a sharper earlier refutation due to Kochen and Specker is not mentioned) and a possibly satisfactory "Prism Model" whose similarity to the so-called "loophole" models and vulnerability to experimental refutation is not mentioned (until the last essay).

Einstein's realism is analysed next, understood as a faith, and dubbed "Motivational Realism". Fine shows that Einstein, while driven by a passionate realism, drew back from making logical assertions connecting physics to that reality. Apparently philosophers are not so judiciously reticent, and Fine rebuts both realist and antirealist positions. Instead he opts for NOA, a minimalist position, that accepts the truth of what is useful (for example the electron as a set of working ideas and experimental results) without further interpretation.

The NOA may be adequate for a spectator of the game of physics, but I do not see it as appropriate for a player. Fine tells us that NOA makes it easy to adopt "paradigm shifts", but that is passive behaviour. Physics, the active game of creating beautiful pictures with predictive power, was played so well by Einstein not least because of his realist faith.

Fine rather snidely compares the realist faith to religious faith. In the same vein I would add philosopher's faith (in unexamined axiomatic concepts), for it seems each philosopher is refuted by another on such "religious" grounds. There is nothing wrong with faith: all physicists act on more knowledge than they can prove. But in matters of faith, certainly evidence counts. The creative successes of realists over the centuries justify their leap of faith more than that of most others. And, although one might attribute the success of quantum theory in the past 60 years to nonrealists, in practice they deal with the reality of the concepts they use the way Bohr dealt with a horseshoe. He put one on his cottage door, and said, "No, I don't believe in it, but I understand the good luck comes even if you don't believe". □

Philip Pearle is a Professor in the Department of Physics, Hamilton College, Clinton, New York 13323, USA.

Making waves on Shelter Island

Ya. B. Zeldovich

Shelter Island II. Proceedings of the 1983 Shelter Island Conference on Quantum Field Theory and the Fundamental Problems of Physics. Edited by Roman Jackiw, Nicola N. Khuri, Steven Weinberg and Edward Witten. MIT Press: 1985. Pp. 369. \$37.50, £37.50.

THE proceedings contained in this book are best viewed against the background of the evolution of physics in the twentieth century and, more specifically, against events at the previous Shelter Island conference of 1947. That meeting was organized by Robert Oppenheimer and was attended by many talented scientists who had returned to fundamental research after working on war projects.

The intermittent character of scientific progress is now well-recognized; bursts of fermentation of new ideas are punctuated by steady elucidation of all the consequences. A Russian poet Tjutchev, practically unknown abroad, said "Blessed are those who come into this world at its fateful moments"; for physicists the remark can be paraphrased as "Lucky are those who attend conferences when fresh ideas are in the air".

One such period was between 1905 and 1928, which witnessed the advent of special and general relativity, quantum mechanics, the quantum theory of fields and the theory of antiparticles. In the 1930s, development slowed down: the refinement of theory was plagued by infinities. A theory of nuclear forces was proposed, with heavy particles — mesons — being an essential part of it. Particles with approximately the appropriate mass were found in cosmic rays, but the euphoria turned to desperation when they were found not to fit the theory. Many eminent physicists believed that a new conceptual framework would be necessary, something as revolutionary as relativity and quantum theory: as Bohr is reputed to have said: "This theory is not crazy enough to be true".

But the younger generation of the time (and many of the old) were more sober. Shelter Island I gave impetus to the work of Schwinger, Feynman and Bethe, whose renormalization programme brilliantly explained the Lamb-Rutherford level shift in the hydrogen atom and the anomalous magnetic moment of the electron. Marshak went on to consider the possibility of two mesons: the muon as a heavy electron and the pion giving nuclear forces. For several decades physics was dominated by ingenious applications of principles formulated before 1930; even

gauge theory and quarks fall in this category. It was not until the mid-1970s and 1980s that a further wave of ideas arose, and it is these new ideas that are the subject of this volume.

In a nutshell, the goal of the old theory was this: given the existence and characteristics of particles and fields, how could one move on to calculate all processes, bound states and so on? The new theory has a more ambitious programme. It aims at explaining *why* all known particles exist — and determining their masses and charges. The same theory also predicts the existence of many as yet unknown particles and fields. Very strange — even "crazy" — ideas are needed!

An excellent overall view of how things stand is given in the first paper in the book, where Murray Gell-Mann provides a coherent yet witty account of how theory has developed between the two Shelter Island conferences. Here, too, Gell-Mann summarizes the topics treated in detail in the following articles — supersymmetry, Kaluza-Klein multidimensional theories, superstrings, and the new and fruitful symbiosis between particle physics and astrophysics.

To a reader not well acquainted with particle physics, I should also add mention of two events of previous history: the introduction of Planck units of mass, length and time in 1899, and the appearance of Dirac's famous paper on large numbers in physics: the ratio of electrostatic to gravitational interaction, 10^{39} in the atom, for example. The modern answer is quite different from that proposed by Dirac, but that does not invalidate the importance of the question he raised.

Characteristic of the new wave of ideas is Steven Weinberg's paper, "Calculation of Fine Structure Constants", in which he discusses the now popular idea of higher dimensional space-time manifolds. Time and three spatial coordinates describe the usual space-time, while the extra spatial coordinates are assumed to be "compactified" to dimensions of the order of 10^{-32} cm. They are not observable as coordinates, but their deformations manifest themselves as fields and particles. As shown in Witten's contribution later in the book, the structure of these extra dimensions may explain the spectrum of fermions — the existence of three, and only three, families of leptons and quarks.

This is a revival of a theory, initiated in the 1920s by Kaluza and Klein, to unify gravity and electromagnetism. For a long time Einstein's own search for a unified theory appeared to be simplistic and to be contradicted by the proliferation of fields and particles found by experimenters. Now, however, most theorists are convinced that the time is ripe for brainstorming in this direction but with more dimensions and a peculiar symmetry

(supersymmetry) between bosons and fermions.

A new branch of physics, as compared with the 1947 conference, is the interdisciplinary study of the very early Universe. Then, because of the negative pressure of a scalar field, one had gravitational repulsion instead of attraction and the radius of the Universe grew exponentially. Both Guth and Linde discuss the topic. (Incidentally, Frenkel and Landau from the USSR were invited to attend Shelter Island I, but neither of them went; Linde, of the Lebedev Institute in Moscow, did attend in 1983, and contributed. The ratio 1:0 is a great improvement!) In the 1960s, the big bang theory of the origin of the Universe was supported by the discovery of background microwave radiation in the Universe, and by the theory of nucleosynthesis occurring at temperatures from 3×10^{10} to 3×10^8 K. These energies are within the range of our particle accelerators, results from which have lent further weight to the big bang theory. But in attempting to go back to an even earlier stage one enters unexplored territory, and the Universe itself has to play the role of the poor man's accelerator. Here the astrophysicist's work is akin to that of the palaeontologist, in which events in a past era must be interpreted from their traces observed today. For example, in my opinion (shared by many others, including Gell-Mann) the very charge asymmetry of the Universe, that is, the absence of bulk antimatter, is proof of the instability of the proton, which Earth-bound experimentalists have searched for in vain.

Everybody working on or interested in the fundamental problems of physics should read *Shelter Island II*. One needs a good grasp of quantum electrodynamics and some special mathematics to appreciate the contents fully, but even those without such knowledge will gain an awareness of contemporary developments in theoretical physics. The science-appeal emanating from the book is very strong; I suspect that many physicists will read it initially without understanding everything, and then learn more fundamentals and return to it.

Although the conference was held in 1983, the book has not dated very much; the various authors remain leading authorities on their subjects and are now, three years later, still developing the ideas discussed in their contributions. Altogether one is grateful to the organizers of the conference and to the editors for bringing together such influential people and for collecting their papers into this excellent book. The publishers might consider issuing a cheaper paperback edition so that its contents would be more widely available, especially to students. □

Ya. B. Zeldovich is at the Institute for Physical Problems, USSR Academy of Sciences, Ul. Kosygina 2, Moscow 117973 GSP-1, USSR.